

ITDL

INTERNATIONAL JOURNAL OF DRAVIDIAN LINGUISTICS

Reg. No. 22651/72

VOLUME XXIII NUMBER 2

JUNE 1994

P.S. Subrahmanyam	<i>Presidential Address to Dravidian Linguists</i>	1
K.S.S. Nambooripad	<i>Hierarchy of Formal Languages</i>	7
A.R. Rajan	<i>Formal Languages Related to Automata</i>	21
T.M. Menon	<i>Joseph's Cat</i>	27
C.A. Winters	<i>The Dravidian and African Languages</i>	34
H. Herring	<i>Great German Dravidologists</i>	53
P. Mohanty	<i>Plural Markers in Kui and Kuvi: A Reconsideration</i>	67
A. Sudharsan	<i>The Impersonal Passive in Kannada</i>	77

Notes

V.S.C. Nair	<i>Case Study of Relativization and Syntactic Markedness of L₂ Learners</i>	87
S. Kaur	<i>Auxiliary Preference and Performance Behaviour of Punjabi Subjects under Contextualized and Non-Contextualized Situations</i>	99
B. Sreedevi	<i>The Language of Administration: Officialese Malayalam - Hindrances in Usage</i>	109
	<i>Silver Jubilee Review</i>	118

Book Review

T.M. Menon	<i>An Anthropological Atlas</i>	130
	<i>Publications Received</i>	140

DRAVIDIAN LINGUISTICS ASSOCIATION

KERALA PAANINI BUILDINGS,
ST. XAVIER'S COLLEGE .P.O.,
TRIVANDRUM - 86, KERALA

EDITORIAL BOARD

Hon. Editor
V.I. Subramoniam

Chief Editor
A.P. Andrewskutty

Deputy Editors

E. James R. Daniel (Nagercoil)
B. Sreedevi (Calicut)
Annie Monsy (Trivandrum)
B. Anvita Abbey (New Delhi)

Editorial Board

James W. Gair (U.S.A)
John F. Mar (U.K.)
A. Govindankutty (Netherlands)
N. Gurov (U.S.S.R.)
K. Kapp (Germany)
E. Annamalai (Mysore)
M. Israel (Madurai)
G. Sambasiva Rao (New Delhi)
B.P. Mahapatra (Calcutta)
K. Rangan (Tanjore)

Committee of Management

President, Vice-President, Secretary and Treasurer of the DLA resident at Trivandrum, Hon. Editor, Chief Editor and Deputy Editors of the *IJDL* resident at Trivandrum, Editor of the DLA News, Director of the ISDL, P. Somasekharan Nair (Trivandrum), J. Nethivanan (Madurai) and Puthusseri Ramachandran.

Endowment Committee

Hans Raj (ICSSR, New Delhi)
G.K. Panikkar (DLA, Trivandrum)
V.I. Subramoniam, Hon. Editor (*Ex-Officio*)

Duration

The journal will be a biannual in English to be published in the first week of January and June each year.

Papers

The manuscripts of articles should be submitted in *triplicate* typed double space with wide margins. Language data should be underlined with meanings in inverted commas. The systems of footnoting and listing of bibliography will be those adopted in *Language*. The article, if theoretically important will be treated as in *Current Anthropology* and published with comments and replies. Fifty offprints will be issued free of cost to the authors. Classical papers which are out of print will also be republished if there is demand.

CHANGE IN DRAVIDIAN GRAMMATICAL SYSTEMS

[Presidential Address to the 22nd All India Conference of Dravidian Linguistics, Trivandrum, 23rd June 1994.]

P.S. SUBRAHMANYAM
Annamalai University

First of all, I wish to express my deep-felt thanks to the members of the Dravidian Linguistics Association for the honour bestowed upon me by electing me the President of the Organisation for the year 1993-94 and thus enabling me to deliver the Presidential Address to the august body of scholars that have assembled here. The International School of Dravidian Linguistics created and nurtured with care by the towering personality of Professor. V.I. Subramoniam and by his younger colleagues here and elsewhere has by now established itself as the only prime institution devoted to the development of Dravidian studies not only in India but also in the entire world. The various standard publications that are brought out so far and those that are under preparation themselves speak for the glory attained by this institution. The scholarly achievements of this institute are responsible in no small measure for the creation of The Indian Institute for Dravidian Studies, which is going to come up shortly at Kuppam (A.P.) with financial assistance from the Central Government and the Governments of the four Dravidian language speaking States. I take this opportunity to welcome all the scholars assembled here to the various forthcoming sessions of this Conference.

I have chosen the topic *Change in Dravidian Grammatical Systems* for this address for the purpose of highlighting the resourcefulness of the Dravidian languages in the matter of expanding their grammatical systems. In the evolution of a language, speakers of later generations often feel the need to represent a distinction between two semantic categories at the morphological level for the sake of clarity. Sometimes, such an urge comes through contact with a neighbouring language in which such a distinction is made overtly. But there are many such cases of development which are due only to the internal mechanism of language change with no discernible outside influence. It is also a fact that all grammatical changes do not contribute to more clarity; the loss of personal suffixes in *Malayālam*, which will be referred to again in the later part of this address, is a well known case which has resulted in ambiguity rather than in clarity.

Change in grammatical systems may be divided into two broad types:

1. Contraction of two independent words into a single word, thus creating a new grammatical category with a new suffix.
2. Analogical creation and extension.

The second process has been recognised as the one mainly responsible for grammar change and many textbooks on historical linguistics cite as examples individual words (or groups of words) affected by it in diverse languages. One of the aims of the present paper is to point out that, in Dravidian, analogy plays a much bigger role in the creation of whole new paradigms by extending the syntactic or morphological distribution of a suffix.

CONTRACTION OF INDEPENDENT WORDS: The present tense in many Dravidian languages takes the cake among the morphological constructions that are formed due to contraction. In Telugu, Old and Modern, it is formed by adding to the present participle of the main verb the forms *unnānu* 'I am', *unnāmu* 'we are' etc., which themselves are contractions of *unna* (past adjective of *uṇḍu* 'to be') + *vāḍanu* (*vāḍu* 'he' + *anu*, a suffix showing agreement with 1st person Sg. pronoun). Thus, for example, Old *cēyucunnānu* (<*cēyucun* + *unna* + *vāḍ-anu*). Modern *cēstunnānu* (<*cēstū* + *unnānu*) 'I am doing', etc.

Koya (a dialect of *Gondi*), on the other hand, retains the original periphrastic construction side by side with the contracted forms, e.g. *ūd-ōr minn-āna* / *ūdōn-āna* 'I am seeing'. *ūd-ōr minn-āma* / *ūd-ōn-āma* 'we are seeing', etc. (-*min*- is deleted in the contracted form; see Subrahmanyam 1971:270).

In Old *Kannada*, the present tense is formed by adding to the past adverbial participle the future tense forms of *āgu* 'to become' e.g. *kēḷ-d* + *a-(p)p-eM* 'I am listening/asking', *kēḷ-d* + *a-(p)p-eM/-ēvu* 'we are listening/asking', etc. (Subrahmanyam 1971:255).

There has been a great deal of discussion on the origin of the present tense in Tamil (summarised in Subrahmanyam 1971:308-10). It is very clear that the finite paradigm with the suffix *-kinr-* / *-kir-* (after a strong verb the *k* of it is doubled) is of late origin and started to appear only in the hymns of the Saivaites and the *Alvārs* (i.e., 7th-8th centuries; see Ramaswami Aiyar 1938:765). Earlier to that, the suffix occurs only in the present adjectival form as in *cēr-kinr-a* 'that which is joining' (*Paripāṭal* 22:35; this is the only instance in Sangam literature according to Ramaswami Aiyar), *ūr-kinr-a* 'that which is spreading', *urul-kinr-a* 'that which is rolling', *ōcaṇi-kkinr-a* 'that which is trying to leave' (these three instances are from *Cilappatikāram*). *Malayālam* *-unn-* is from *- *kinr-*.

Whatever be the origin of this suffix, it is certain that the full finite paradigms with it are made analogically by tagging on the personal suffixes to the form with *-kinr-/-kir-* the model for the analogy being the past and the future paradigms. Andronov (1961) gave two suggestions regarding the origin of this suffix. One is that it is in origin the verb *kil-* (*kirr-*) 'to be able' DEDR 1570; note the Toda cognate *kīs-* (< *kil-*) with the secondary stem *kīd-* (< **kinr-*, which is formally identical with the Tamil suffix under discussion, as has been pointed out by Emeneau (forthcoming, p.12)). Zvelebil (1971) also arrives at the same conclusion. The second suggestion is that the *k* in *cey-kinr-ēn* 'I am doing' etc., is the non-past suffix *k/kk* and that *i-nrēn* is the past form of the verb *il*, which does not occur otherwise. The present author (1971:309) meekly suggested that the verb involved here could be the negative verb *il*, but this is highly improbable since *il*-nowhere else occurs in the affirmative. The first suggestion of Andronov is supported in recent times by Steever (1989) (and Emeneau forthcoming, expresses his approval for it) though neither of them have given any satisfactory explanation for the chief problem raised by the present author (1971:309): *nr* in *-kinr-* as opposed to *rr* in the past stem *kirr-* of *kil-*. Emeneau, however, notes that the present of the Toda verb with the required past stem *kīd-* < **kinr-* eases the problem and that even in Tamil some verbs ending in *l* have doublets in the past stem with *nr/rr* (e.g. *ēl* 'to happen' (*ēr-ēr-*), DEDR 905, which makes the existence of unattested *kinr* a possibility. Thus, even this explanation is not completely free from problems. We must leave this problem at this for the present, noting, however, that *-kir-* could be a weakened form of *kirr-* (cf. for similar weakening, *irupatu* '20' < **irupattu*, *muppatu* '30' < **muppattu* etc.) since it could not be directly from *kinr-* as is generally supposed because of the absence of forms showing a similar loss of the nasal elsewhere (in spoken Tamil **nr* > *nn/nn*). The occurrence of the *-kir-* form in *Maṇimēkalai* itself (*cātikkira* 'that which achieves' 29:298 besides *enkinra* 'that which says' 29:293) supports the conclusion that *-kir-* cannot be from *kinr-* by the loss of the nasal.

The Toda past tense with *-š-* as the past suffix at the surface level has been explained as a periphrastic construction in origin with the past adverbial form of the main verb + the non-past forms of **iru* 'to be' (Subrahmanyam 1991:62ff). The past tense paradigm of *po* :- (*po-d-*) 'to come' from **va* :- (*va-nt-*) are given below:

Sg.	Pl.
1. <i>podšpen</i>	(Excl.) <i>podšpem</i> (Incl.) <i>podšpum</i>
2. <i>pošpy</i>	<i>podš</i>
3. <i>podč</i> (< <i>podšt</i>)	<i>podč</i> (< <i>podšt</i>)

Emeneau earlier suggested the possibility of connecting this -š- suffix to the reconstructed past suffix *-c- found in Old Ta. (e.g. *va-nt-icin* 'I came', etc.) and other languages. But such a connection is fraught with problems since the correspondence *c: To. š is otherwise unattested and since this does not explain the additional consonant found in the paradigm. Treating these forms as contractions avoids these problems without giving rise to any other new problem. For example, *podšpen* is from **va-nt(u)* + **iru-ppēn* 'I have come/came' (with *r* > To. š, which is well-attested and **pp* > *p*).

Instances of contraction outside the tense system are found in the obligative, the permissive and the imperative forms of some of the languages. Examples: Ta. *var-alām* 'one may come' < *var-al* 'coming' + *ākūm* 'will/may happen', (spoken) *ceyy-aṇum* 'one must do' < *ceyy-a* + *vēnt-um* 'lit., doing is necessary', Tel. (Mod.) *ceyyāli* 'one must do' < (Old) *cēyā* + *valayunu* with the same structure and meaning as the above construction in Tamil. *Koḍaḡu keyy-onḍu* 'one must do' < **keyy-a* + *vēntā*, negative verb from *vēntu* 'to be required'. *Tulu pōv-oḍu* 'one must go' < **pō(k-a)* + *vēnt-um*, *bar-aḍ-a* (Sg.)/*bar-aḍ-e* (Pl.) 'do not come' < **var-a* + *vēntā* (with the Sg./Pl. distinction being a later creation on the analogy of the corresponding affirmative forms, e.g. *koru* (Sg.)/ *kor-i* (Pl.) 'Give!').

ANALOGICAL CREATION AND EXTENSION: Analogical creation and extension is well-documented in the Dravidian family. The close-knit group of personal pronouns provides numerous instances for it (see Subrahmanyam 1993). The oblique base of these pronouns, being statistically more predominant than the nominative base, either forces the latter to take an additional initial consonant to bring about more similarity (e.g. Tel. *ēnu* > *nēnu*) 'I', *ēmu* > *mēmu* 'we (Excl.)' under the influence of the oblique bases *nā-* and *mā* respectively or, more drastically, replace the latter (e.g. Ma. *ñāñal* > *ñanñal* 'we (Excl.)', Toda *em* 'we (Excl.)' < oblique **yam-*, *om* 'we (Incl.)' < oblique **nam-*). In *Gondi* was created a new 2nd person Sg. form *immē/nimmē* 'you (Sg.)', which is not directly relatable to P. Dr. **nīn*, by cutting off the Pl. suffix in the corresponding Pl. form *immat/nimmat* 'you (Pl.)'. In *Gondi*, the 3rd non-human Pl. suffix in finite verbs comes from the plural suffix *-n* that occurs in non-human nouns ending in a vowel (e.g. *kāyā- n man-t-ān* (unripe) fruits are (there)', *paddi- n vā-t-ān* (pigs came)' as has been noted long ago by Jules Bloch (see Subrahmanyam 1971:420). *Pengo -in* also has the same history. This suggests that originally in *Gondi* (and in *Pengo*) the 3rd non-human Sg. form was used in agreement with plural nouns as is the case in some of the Dravidian languages.

The influence of analogy is seen on a much wider scale in the creation of new verbal paradigms in a number of languages. In *Malayālam*, as shown by Ramaswami Aiyar, there has been a gradual replacement of verb forms with personal suffixes by those without. However, it is to be emphasised that this is not a pure phonological change since no similar development is found elsewhere in the language. Analogical levelling has been at work here. This is evident from the future tense forms with the suffix *-um*. Old Tamil allows the *-um* forms to agree with *avan* 'he', *aval* 'she', *atu* 'it' and *avai* 'they (neut.)'. *Malayālam* went one step ahead by allowing them to agree with the pronoun of any person or number, e.g. *ñānceyyum* 'I will do' *avar ceyyum* 'they will do' etc. This pattern served as the analogy for the past and the present forms to shed their personal suffixes so that forms like *ñān cey-tu* 'I did', *ñān ceyy-unnu* 'I am doing' (*-unnu* < **-kinr-*) became firmly established.

Even in Old Tamil itself, *avan ceyyum*, *aval ceyyum* are replaced later by *avan cey-v-ān* and *aval cey-v-āl*, which contain personal suffixes, for the sake of avoiding ambiguity. The past tense (e.g. *cey-t-ān*, *cey-t-āl*) and the present tense forms (e.g. *cey-kinr-ān*, *cey-kinr-āl*) must have served as the model for this. Similar analogical developments can be seen in Telugu also. Corresponding to Old Telugu *vāḍu/adi/avi cēs-unu* 'he/she/it will do' we have in Modern Telugu *vāḍu/cēs-tāḍu*, *adicēs-tun-di* and *avi cēs-tāyi*; cf also the past tense forms: (Old) *cāḍu/adi/avi cēy-enu* 'he/she/it did' and (Mod.) *vāḍu cēs-ē-ḍu*, *adi cēs-in-di* and *avi cēs--ē-yi*. In both Tamil and Telugu, the newer forms contain tense suffixes that are different from those found in the older forms.

Kota affords an obvious instance of the creation of an entire paradigm on the analogy of another paradigm. The past tense paradigm of this language acquired the additional consonant *v* on the analogy of the future tense paradigm in which the consonant is inherited from the earlier stages as is evident from the following paradigms of the verb *vār-(va-d-)* 'to come' (for further details, see Subrahmanyam 1991:58). Note that the past forms of other closely-related languages have no *v* in them, for example, Ta *va-nt-ēn*, Ka. *ba-nd-enu* etc.

	Past	Future
I P. Sg.	<i>vadvē</i>	<i>vakvē</i> (< <i>*varu-kuv-en</i>)
I P. Pl.(Excl.)	<i>vadvēm</i>	<i>vakvēm</i>
I P. Pl.(Incl.)	<i>vadvōm</i>	<i>vakvōm</i>
II P. Sg.	<i>vadvī</i>	<i>vakvīm</i>
II P. Pl.	<i>vadvīm</i>	<i>vakvīm</i>
III P.	<i>vadukō</i>	<i>vakugō</i>

The creation of the present tense paradigm in Parji is unique. It has restricted the non-past suffix **-um* (> Pa.-m-) to the present tense and

added the regular personal suffixes after it to form a full paradigm (Subrahmanyam 1964). The paradigm from *ver-* 'to come' is given as an example below:

	Sg.	Pl.
I P.	<i>ver-m-en</i>	<i>ver-m-om</i>
II P.	<i>ver-m-ot</i>	<i>ver-m-or</i>
III P. Mas.	<i>ver-m-ed</i>	<i>ver-m-er</i>
III P. Non-Mas.	<i>ver-m-o</i>	<i>ver-m-ov</i>

Note that the present stem here (i.e. *ver-m-*) is identical with the future form *varum* of *Tamil-Malayalam*; the personal suffixes in this language were added analogically (on the basis of the past paradigm). Parji verbs that correspond to the strong verbs of the Southern languages have the formative *-k* in the Northern dialect (but it is *-p* in the Southern dialect before this) just as *kk-* appears before *-um* in those verbs of the Southern languages, e.g. (N.) *nil-k-m-en*/(S.) *nil-p-m-en* 'I am standing': Ta. *nirk-um*, Ma. *nilkk-um* (Subrahmanyam 1971:312). This is further evidence for connecting the Parji *-m-* with original non-past **-um*. The *p* of such forms in the Southern dialect is the analogical replacement of *k*, which is the form that is historically expected in this context, although the *p* (< **non-past*-pp-*) is also a non-past suffix as old as **-kk-*.

Bibliography

- Andronov, M. 1961. "Hints regarding the origin of the present suffix 'kinr' in Tamil". *Tamil Culture* 9. 145-50.
- Emeneau, M.B. (forthcoming). *The palatalizing rule in Tamil-Malayalam and Telugu*.
- Ramaswami Aiyar, L.V. 1938. "The morphology of the Old Tamil verb". *Anthropos* 33. 47-81.
- Steever, Sanford B. 1989. "On the etymology of the present tense in Tamil". *JAOS* 109. 237-54.
- Subrahmanyam, P.S. 1964. "Two problems in Parji verb forms". *Indian Linguistics* 25. 47-55.
-1971. *Dravidian verb morphology - a comparative study*. Annamalainagar: Annamalai University.
-1991. "Tense formation in Kota-Toda : a comparative study". *Studies in Dravidian and General Linguistics (a Festschrift for Bh. Krishnamurti)*: 49-72.
-1993. "The personal pronouns in Dravidian". *PILC Journal of Dravidic Studies* 3 (no.1): 27-52.
- Zvelebil, Kamil.V. 1971. "The present tense morph in Tamil". *JAOS* 91: 442-5.

HIERARCHY OF FORMAL LANGUAGES

K. S. S. NAMBOORIPAD

Department of Mathematics, University of Kerala

June 21, 1994

ABSTRACT: In this talk my aim is to give a brief introduction of the Mathematical Theory of Formal Languages. We discuss the concepts of syntactic equivalence of words in formal languages, recursive and recursively enumerable languages, Turing machines and Markov's normal algorithms. We then introduce the concept of formal grammars, which leads to a classification of languages in terms of the grammars generating them. We end the talk with a definition of Chomsky hierarchy of formal languages.

• Introduction

A language is a communication devise. It consists of *words or sentences* while a word itself is an array of letters in the *alphabet*. It is clear that, given an alphabet, not all arrays will be acceptable as words. The principal question dealt with in Mathematical linguistics is the following: Given a language over certain alphabet, how to decide when a given array constitute an acceptable word in the language ?

There are two approaches to this problem. We may say that a word is meaningful if it conveys some 'meaning'. However, this is not a deterministic phenomenon as the meaning depends on many factors. The study of this aspect is known as *semantics or communication theory* and, in mathematical formulations, use the frame work of probability theory (see [2,7]). On the other hand, the study of the structure of words itself is called its *syntax* and the theory which studies this aspect was called by Chomsky as *algebraic linguistics* (see [3] as well as [8]). This is a deterministic theory and the frame work it uses is that of logic and semigroup theory. In this talk I shall attempt to give an over view of this latter theory which is at present a significant part of the combinatorial semigroup theory. We refer the reader to [1,8,9,12] for details for various topics we have discussed below.

• Preliminaries

We begin by recalling some of the mathematical concepts which we will have to use constantly in the sequel. Informally, a *set* is a collection of 'objects'. Symbols A , X , etc. will be used to denote sets and if a is an object in the set A , we write $a \in A$ (a is a member of A). Here \in is called the *membership relation* and is a two-place predicate. We use the usual methods for specifying sets: A set A is specified either as

$$A = \{a_1, a_2, \dots, a_n\} \quad \text{or} \quad A = \{x : P(x)\}.$$

The first method is used if the complete list of objects in A is known; in this case, the set is specified by giving the complete list of (names of) object enclosed in braces. In case it is not possible to give a complete list of objects of A but we know some property common to all objects of A , then the second method is used. Notice that in the second specification, $P(x)$ is actually a one-place predicate which becomes a true statement when we replace the variable x by (the name of) an object in A and false when the object is not in A . In this case, the set A is called the *extensionality* of the predicate P . This principle which asserts that 'extensionalities of certain predicates are sets', is called the *Axiom of specification* [The reader is warned that unrestricted use of this principle leads to contradictions. Thus use of the predicate $P(x) = 'x \text{ is not a member of } x'$ leads to the paradox known as *Russel paradox*—see [10,21] for a more detailed discussions of the axiom of specification and how its improper use leads to the Russel paradox. This is partially the reason why we have only *described* and not *defined* sets above.]

Elementary operations on set can be defined in terms of logical connectives using this specification. Thus intersection, union, inclusion, equality and difference of two sets $A = \{x : P(x)\}$ and $B = \{x : Q(x)\}$ are defined in terms of the basic *logical connectives* conjunction (in symbol \wedge), disjunction (\vee), implication (\implies), double implication (\iff) and negation (\neg) as:

$$\begin{aligned} A \cap B &= \{x : P(x) \wedge Q(x)\}, & A \cup B &= \{x : P(x) \vee Q(x)\}, \\ A \subseteq B &\iff (P(x) \implies Q(x)), & A = B &\iff (P(x) \iff Q(x)) \\ & & &\text{and} \\ A \setminus B &= \{x : P(x) \wedge \neg Q(x)\}. \end{aligned}$$

If a_1, a_2, \dots, a_n are objects, we denote by (a_1, a_2, \dots, a_n) the object called the *ordered n -tuple* of objects a_i (ordered pair if $n = 2$, ordered triple if $n = 3$ and so on); here a_i is called the *i -th co-ordinate* of the new object (a_1, a_2, \dots, a_n) for $i = 1, 2, \dots, n$ (a_1 being the first, a_2 , the second and so on). If A_1, A_2, \dots, A_n are sets, we set

$$\prod_{i=1}^n A_i = A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) : a_i \in A_i, i = 1, 2, \dots, n\}$$

The set $\prod_{i=1}^n A_i$ is called the *Cartesian product* of the sets A_1, \dots, A_n . If all the sets A_i are equal to (say) A , then the Cartesian product of these sets is denoted by A^n .

A subset R of $A \times B$ (Cartesian product of two sets) is called a *relation* from A to B (or a relation between objects of A with that of B). In dealing with relations, it is traditional to indicate membership $(a, b) \in R$ (in a given relation $R \subseteq A \times B$) as aRb which suggest that R is (logically) a two-place predicate. In fact any two-place predicate xPy (where x and y are logical variables) its extensionality is a relation as defined above provided that the extensionality is a set. In view of the definition of equality of sets, the two approaches are equivalent.

Given a relation $R \subseteq A \times B$, we set

$$\text{dom } R = \{a : a \in A, \text{ and there exists } b \in B \text{ with } (a, b) \in R\}$$

$$\text{Im } R = \{b : b \in B, \text{ and there exists } a \in A \text{ with } (a, b) \in R\}$$

Note that $\text{dom } R$ is a subset of A called the *domain* of R and $\text{Im } R$ is a subset of B called the *range* (or image) of R . R is said to be *single-valued* if it satisfies the condition that for each $a \in \text{dom } R$ there is a *unique* $b \in \text{Im } R$ such that $(a, b) \in R$. The uniqueness of the second coordinate (given the first), implies that for each $a \in \text{dom } R$, we may unambiguously represent the second coordinate as $R(a)$ (some writers use the notation aR). In this case R is also called a *partial function* of A into B (written as $R: A \rightarrow B$). In case we also have $\text{dom } R = A$ we say that R is a *function*. We shall use symbols f, g , etc. to denote functions or partial functions. Thus a relation $f \subseteq A \times B$ is a function $f: A \rightarrow B$ if

$$(a, b), (a, b') \in f \implies b = b' \quad \text{and} \quad \text{dom } f = A.$$

Informally, we may think of a function $f: A \rightarrow B$ as a process which when applied to an element $a \in A$ produces the unique element $f(a) \in B$ such that $(a, f(a)) \in f$. $f(a)$ is called the *value* of f at $a \in A$. A function (or partial function) $f: A \rightarrow B$ is *one-to-one* or *injective* if $f(a) = f(b)$ implies $a = b$ (so that distinct elements of A generate distinct elements of B under the process f). It is said to be *onto* or *surjective* if $\text{Im } f = B$. f is said to be *bijective* if f is both injective and surjective; in this case, the set A is said to have a *one-to-one correspondance* with B . A *partial one-to-one correspondance* of A with B is a one-to-one correspondance of a subset of A with B .

As an example, consider the set $\mathbf{N} = \{1, 2, \dots\}$ of all natural numbers. We denote by \mathbf{N}^0 , the set $\mathbf{N} \cup \{0\} = \{0, 1, 2, \dots\}$. The subset $\mathbf{Nr} = \{1, 2, \dots, r\}$ is called an initial segment of \mathbf{N} determined by $r \in \mathbf{N}$ and $\mathbf{Nr}^0 = \mathbf{Nr} \cup \{0\}$ is an initial segment of \mathbf{N}^0 . A *counting* of a set A is a one-to-one correspondance whose domain is either \mathbf{N} or \mathbf{Nr} for some $r \in \mathbf{N}$ with range A . Of course if there is a counting of A

with domain $\mathbf{N}r$, then A has exactly r elements; such sets are said to be *finite*. Otherwise, A is said to be *countably infinite*. There are sets that does not admit any counting.

Let $f: A \rightarrow B$ be a function. If $A \subseteq A_1 \times \cdots \times A_n$ then f is said to be a *partial function of n variables*. The value of f at $(a_1, \dots, a_n) \in A$ is denoted by $f(a_1, \dots, a_n)$. f is a *function of n variable* if $A = A_1 \times \cdots \times A_n$. A partial function $f: (\mathbf{N}^0)^n \rightarrow \mathbf{N}^0$ is called an *integral function of n variable*.

If $R \subseteq A \times B$ is any relation, then $R^{-1} = \{(b, a) : (a, b) \in R\}$ is a relation from B to A ; it is called the *inverse* of R . If $R \subseteq A \times B$ and $S \subseteq B \times C$ are relations we can form a new relation from A to C as follows:

$$R \circ S = \{(a, c) \in A \times C : \text{for some } b \in B, (a, b) \in R, \text{ and } (b, c) \in S\}$$

This relation is called the *composite* of R and S . A relation $R \subseteq A \times A$ is called a relation in the set A . We use the notation 1_A to denote the relation $\{(a, a) : a \in A\}$; it is called the *identity relation* in A . A relation R in A is said to an *equivalence relation* if it has the following properties:

- *Reflexive*: $1_A \subseteq R$;
- *Symmetric*: $R^{-1} \subseteq R$; and
- *Transitive*: $R \circ R \subseteq R$.

If R is an equivalence relation in A , then for $a \in A$ the set $aR = \{a' \in A : (a, a') \in R\} \subseteq A$ is called the *equivalence class* of a . Note that in the definition of the equivalence relation implies that for $a, b \in A$, either $aR = bR$ or $aR \cap bR = \emptyset$ (here \emptyset denote the set having no element so that the last condition implies that aR and bR have no element in common—such set are said to be *disjoint*). In view of the first condition above, every $a \in A$ belongs to some equivalence class so that $\cup_{a \in A} aR = A$. We say that the equivalence relation R induces the *partition* $\{aR : a \in A\}$ of A .

• Languages

In the context of language theory, a finite sets is often called an *alphabet*. Let $A = \{a_1, a_2, \dots, a_n\}$ be an alphabet (finite set). A finite sequence $\{a_{i_1}, \dots, a_{i_r}\}$ of elements of A is called a *word* over the alphabet A and is indicated by writing the (names of) elements as an array; thus $a_{i_1} \dots a_{i_r}$ is the word determined by the sequence $\{a_{i_1}, \dots, a_{i_r}\}$. We denote by A^* , the set of all words over A . In A^* , we also include the word 1 having no symbol (corresponding to empty sequence). If $w = x_1 \dots x_n$ and $w' = y_1 \dots y_m$ are words over A , we form a new word ww' as follows:

$$ww' = x_1 \dots x_n y_1 \dots y_m.$$

This gives a *binary operation* on the set A^* of all words in the alphabet A which is called the *concatenation or product* of w and w' . This operation is an *associative* on A^* ; that is, given $u, v, w \in A^*$, we have

$$(uv)w = u(vw). \text{ Also, clearly, } w1 = 1w = w$$

for all $w \in A^*$. This means that 1 is the identity element for the product of words in A^* . Recall that a set with an associative binary operation is called a *semigroup* and that a semigroup is a *monoid* if the binary operation admits (a unique) *identity*. Thus A^* is a *semigroup with identity* or a *monoid* (see [12]). A^* is called the *free monoid* on A . The subset $A^+ = A^* - \{1\}$ consisting of all non-empty words is closed with respect to the product defined above and so A^+ is a semigroup. A^+ is called the *free semigroup* on A (A^+ has no identity).

- A language over the alphabet A is a subset of A^* .

To illustrate this definition, we give a few examples.

EXAMPLE 1.

Let $A = \{\mid\}$ (the singleton set having only "stick" as element). A^* consists of empty array and all arrays of the form $\mid, \mid\mid, \mid\mid\mid, \dots$. Any sub-collection of these is a language over L .

EXAMPLE 2.

Let $A = \{ (,) \}$ (opening and closing of brackets). A^* contains arrays such as $(,), (, (,)(,)(, \dots$. A language over A is obtained by requiring that the arrays must consist of balanced (correct) brackets such as $(,) (,)(, ((,))$, etc. but not $),)(,)(,)(, \dots$. This language is called the *restricted Dyke language* on two letters or the *correct parenthesis language*. Though quite simple, it is nevertheless useful. An implementation of this language is included in many word-processor programs.

EXAMPLE 3.

Let A denote the set of roman letters. Then the set of all words and sentences in English is a language over A . On the other hand, English language can be regarded as the set of all correct sentences. Hence if B denote the set of all English words, then English Language is a language over B in the sense defined above.

The process of expressing one alphabet as words in another used in the example above is of interest in severable occasions. For example, ASCII symbols, numbers, etc. are represented in the central processors of computers by using appropriate codes in bits (alphabet $\{0, 1\}$).

• Syntactic congruences and the word problem

Let $L \subseteq A^*$ be a language over the alphabet A . If $w \in A^*$, the *context* of w with respect to L is defined as

$$\text{Cont}_L(w) = \{(u, v) : u, v \in A^* \text{ and } uwv \in L\}.$$

It is clear that $\text{Cont}_L(w) \subseteq A^* \times A^*$ and hence it is a relation in A^* . We now set

$$P_L = \{(w, w') : \text{Cont}_L(w) = \text{Cont}_L(w')\}.$$

It is clear that P_L is an equivalence relation (this is a simple consequence of the logical properties of the *equality* relation. In addition, P_L has the following property:

$$(w, w') \in P_L \implies (uwv, uw'v) \in P_L$$

for all $u, v \in A^*$. Equivalence relations with this property are called *congruences*. When $\text{Cont}_L(w) = \text{Cont}_L(w')$, we say that w and w' are *syntactically equivalent* and P_L is called the *syntactic congruence* of the language L . Now P_L divide A^* into disjoint equivalence classes $\{wP_L : w \in A^*\}$ and each element of this new set may be thought of as a *syntactic element* with respect to L . We denote the set of syntactic elements by M_L . The map $P_L^h : A^* \rightarrow M_L$ which sends $w \in A^*$ to the element $wP_L \in M_L$ has the property that

$$P_L^h(w)P_L^h(w') = P_L^h(ww').$$

Thus we can define a multiplication in M_L by $(wP_L)(w'P_L) = ww'P_L$ which is associative and the map P_L^h preserves the multiplication (in the sense that the image under P_L^h of a product ww' in A^* is the product of images). Thus M_L is a monoid and P_L^h is a homomorphism. M_L is called the *syntactic monoid* of L and P_L^h is called the *syntactic homomorphism*.

One of the important questions in the formal theory of languages is to determine when two words w, w' in the language L has the same syntax. By the above, this is equivalent to determining whether or not wP_Lw' . Note that a direct use of the definition of the congruence P_L is not possible in most cases. This leads us to the following problem, first formulated by Thue [22] for semigroups. We state this in the form suitable for language theory. Here, by an *algorithm*, we mean a process consisting of a finite number of explicit steps (which, in theory, can be performed by a machine with out further human intervention).

- WORD PROBLEM FOR LANGUAGES: Let L be a language over a finite alphabet A . Given $w, w' \in L$ (or, more generally, $w, w' \in A^*$)

A^*), does there exists an algorithm to determine whether or not w and w' have the same syntax with respect to L ?

This problem was solved negatively by logicians E. L. Post [19] and A. A. Markov [14,15] in 1947. This implies that, in general, it is not possible to settle the question algorithmically whether two given words of a language has the same syntax or not. Another problem which is important in the theory of formal languages is the following:

- ACCEPTOR PROBLEM: Given a language L over the finite alphabet A , does there exists an algorithm to decide whether an arbitrary word $w \in A^*$ belongs to L or not.

Again, the situation is similar to the word problem: in general, there is no algorithm to decide whether $w \in L$ or not.

It is clear that, to get better information about these problems, one has to consider special classes of languages.

• Recursively enumerable languages

In order to investigate classes of languages for which there are general methods for dealing with problems mentioned (and other similar problems), three equivalent logical concepts have been introduced. These are:

- (1) *partial recursive functions* by S. C. Kleene (1936—see [11]);
- (2) *Turing machines* by A. M. Turing (1937—see [23]); and
- (3) *normal algorithms* by A. A. Markov (1956—see [17]).

We shall give a brief description of these.

Partial recursive functions: A partial function $f: (N^0)^n \rightarrow N^0$ is said to be partially recursive if it can be built up from the basic functions

$$suc: x \mapsto x + 1, \quad zr: x \mapsto 0, \quad p_i^n: (x_1, \dots, x_n) \mapsto x_i$$

by means of the following procedures:

- (i) composition, (ii) primitive recursion, and (iii) minimalization

We shall not go into the details of these; instead we shall give an example to illustrate the use of these.

EXAMPLE 4.

The sum-function $s: (N^0)^2 \rightarrow N^0$, defined by $s(x_1, x_2) = x_1 + x_2$ is partially recursive. For let h denote the composite of p_1^2 and suc so that

$$h(x_1, x_2) = suc(p_1^2(x_1, x_2)) = x_1 + 1.$$

Then define s by the following primitive recursion:

$$\begin{aligned} s(x_1, 0) &= x_1; \\ s(x_1, y+1) &= s(h(x_1, x_2), y) \end{aligned}$$

Similarly, the function $d(x_1, x_2) = x_1 - x_2$, $x_1 \geq x_2$ is also partially recursive. For we can define this in terms of the sum-function s using the minimalization as follows:

$$d(x_1, x_2) = \mu_y [s(x_2, y) = x_1]$$

where the right-hand side denotes the minimum solution in y of the equation $s(x_2, y) = x_1$.

A set $Y \subseteq N^0$ is said to be *recursive* if the function defined by

$$\chi_Y(y) = \begin{cases} 1, & \text{if } y \in Y \\ 0, & \text{if } y \in N^0 \setminus Y \end{cases}$$

is partially recursive. χ_Y is called the characteristic function of Y . Y is said to be *recursively enumerable* if there is a partial recursive function f such that $Y = \text{Im } f$. Any recursive set is recursively enumerable; but the converse is not true. There exist examples of recursively enumerable sets which are not recursive (see [12]).

Let $A = \{a_1, a_2, \dots, a_n\}$ be a finite alphabet. We can define a *numbering* of A^* as follows: if $w = a_{i_0} \dots a_{i_k} \dots a_{i_r}$ (where $1 \leq i_1 \leq \dots \leq i_k \leq \dots \leq i_r \leq n$) is any word in A^* , let

$$\lambda(w) = i_0 + i_1 n + \dots + i_k n^k + \dots + i_r n^r$$

It can be checked that $\lambda: A^* \rightarrow N$ is a bijection. Using this we can convert any partial word function $f: (A^*)^m \rightarrow A^*$ to a partial integral function. For, if x_i , $i = 1, 2, \dots, m \in N^0$, replace them by the above, corresponding to each integer x_i there exists unique word $w_i \in A^*$ such that $x_i = \lambda(w_i)$ for $i = 1, 2, \dots, m$. Then we set

$$\lambda(f)(x_1, \dots, x_m) = \lambda(f(w_1, \dots, w_m)).$$

We say that the partial word function f is partially recursive if $\lambda(f)$ is partially recursive as defined above. Also, a language $L \subseteq A^*$ is recursive (or recursively enumerable) if the subset $\lambda(L) = \{\lambda(w) : w \in L\}$ is a recursive (respectively, recursively enumerable) subset of N^0 . We refer the reader to [13] for details.

Turing machines: A *Turing machine* T is a quadruple (Q, A, b, π) where

- (a) Q is a finite set called the *set of states* of T ; we adjoin a distinguished symbol s , called the *stop symbol*, to Q and we denote the new set $Q \cup \{s\}$ by Q'
- (b) A is a finite set disjoint from Q' and is called the alphabet of T ; again we adjoin two new symbols r and l called as the right symbol and the left symbol respectively. We let $A' = A \cup \{r, l\}$.
- (c) b is an element of A , called the *blank symbol*.
- (d) $\pi: Q \times A \rightarrow Q' \times A'$ is a function, called the program of T .

Intuitively, Turing machine can be described as a device having along input-output tape divided into cells and a read-write head which scans the cells one by one from left to right. At any given instant suppose that the machine is in the state q and scanning the cell containing the letter a . If $\pi(q, a) = (q', a')$, the machine change the state to q' and replace the letter a by a' in the cell. Suppose that initially the machine is in the state q and that the tape carries the word $w = u_1 \dots u_r$ where $u_i \in A$. After some computation suppose that the machine has reached the state q_i and scanning the letter u_k . At this point, the machine can perform the following operations (depending on the programme π):

- (a) Change the state from q_i to q_m and replace u_k by a new letter $v_k \in A$ so that $\pi(q_i, u_k) = (q_m, v_k)$.
- (b) Change the state from q_i to q_m and moves to the next cell on the right. In this case, we have, $\pi(q_i, u_k) = (q_m, r)$.
- (c) Change the state from q_i to q_m and erases the letter in the cell left of the cell containing u_k . In this case, we have, $\pi(q_i, u_k) = (q_m, l)$.

The machine stops when it reaches the state s . The word w' which remains on the type at the time when it reaches the state s is the word computed by the machine from the word w on the tape when the machine is started with the initial state q . We write $w' = T_q(w)$. Let f be a partial word function of m variables over a finite alphabet A . We say that f is Turing computable if there exists a Turing machine $T = (Q, A', b, \pi)$ with $A' = A \cup \{b\}$ and a state $q \in Q$ such that for all $w_i \in A^*$, $i = 1, \dots, m$, we have

$$f(w_1, \dots, w_m) = T_q(\bar{w})$$

where $\bar{w} = bw_1bw_2 \dots bw_m$ is the word obtained from the words w_i , $i = 1, \dots, m$ by introducing blank symbols and forming the product. The last equation means that the left-hand side is defined if and only if the word $\bar{w} = bw_1bw_2 \dots bw_m$ is Turing computable and that the computed result is the value of the function. We have the following result (see [13]):

- *The class of all partial recursive functions coincides with the class of all Turing computable functions.*

This result shows that the theory of recursive and recursively enumerable sets coincides with that of Turing computation.

Normal algorithms: Let A be a finite alphabet and let $A' = A \cup \{\rightarrow, \cdot\}$ where \rightarrow and \cdot are symbols not in A . A *normal algorithm* (or just algorithm) \mathcal{A} is a finite ordered sequence of *substitution formulas* of the following type:

$$(1) : u \rightarrow v \quad \text{or} \quad (2) : u \rightarrow \cdot v$$

where $u, v \in A^*$. Note that the formulas are just words in A' with a left side u and a right side v belonging to A^* . Given $w \in A^*$, we define the *direct transform* of w as follows:

Case 1: No left side of the substitution formulas is a factor of w . In this case, we say that \mathcal{A} does not apply to w .

Case 2: At least one left side of the substitution formula is a factor of w . Let F_1 denote the *first formula* in the list whose left side u_0 is a factor of w . Suppose that $w = w_1 u_0 w_2$ so that u_0 is not a factor of w_1 . Then we replace w by $w' = w_1 v_0 w_2$ if F_1 is of type (1) and by $\cdot w'$ if F_1 is of type (2). This is indicated by writing $w \rightarrow w'$ in the first case and $w \rightarrow \cdot w'$ in the second case; in both case, we say that w' is a direct transform of w .

We write $\mathcal{A}(w) = w'$ to mean that either \mathcal{A} does not apply to w and $w = w'$ or there exists a finite sequence $w_i : 0 \leq i \leq s$ such that $w_0 = w$, $w_s = w'$ and w_{i+1} is a direct transform of w_i , $0 \leq i \leq s-1$. A partial word function f in m variables over the alphabet A is *normally computable* if there exists a normal algorithm over the alphabet $A' = A \cup \{b\}$ (where $b \notin A$) such that

$$f(w_1, \dots, w_m) = \mathcal{A}(bw_1bw_2 \dots bw_m)$$

for all $w_i \in A^*$, $i = 1, 2, \dots, m$. Again we have

- The class of all normally computable partial word functions coincides with the class of all partial recursive word functions.

The foregoing discussions shows that the concepts of *recursiveness*, *Turing computability* and *normal computability* are logically equivalent concepts which is useful in studying formal languages.

• Formal Grammars

For languages, there is yet another useful formalism, introduced by N. Chomsky and others, which is equivalent to those discussed above. This is an extreme formalization of the traditional grammatical analysis of natural languages.

A *formal grammar* Γ is a triplet (V, A, π) where

- (a) V is a finite set, called the *vocabulary* of Γ ;
- (b) A is a subset of V , called the *terminal alphabet* of Γ and $B = V \setminus A$ is called the set of *auxiliary symbols* of Γ ; and
- (c) $\pi \subseteq B^+ \times V^*$ is a relation from B^+ to V^* (here B^+ denote the set of all non-empty words in B). Ordered pairs belonging π are called *productions* of Γ , and we write $u \rightarrow v$ if $(u, v) \in \pi$.

A word $w' \in V^*$ *derives directly* from a word w (written $w \rightarrow w'$) if there exist words $x, y \in V^*$ and a production $u \rightarrow v$ such that $w = xuy$ and $w' = xvy$. We say that a word w' *derives* from w (written $w \xRightarrow{\Gamma} w'$) if either $w = w'$ or there exists a finite sequence w_0, \dots, w_s such that

$$w = w_0 \rightarrow w_1 \rightarrow \dots \rightarrow w_s = w'.$$

For every $\sigma \in B$, the set

$$L(\Gamma, \sigma) = \{w \in A^* : \sigma \xRightarrow{\Gamma} w\}$$

is called the *language generated by Γ from σ* . A language L over the alphabet A is a language generated by Γ if $L = L(\Gamma, \sigma)$ for some $\sigma \in B$.

For example, if Γ represents a formalization of the usual English grammar, then, we may assume that the set of auxiliary symbols B will include all grammatical terms like (sentence), (noun), (article), (verb), etc. and the terminal alphabet will include all English words. Also the set of production π must contain rules such as:

$$\begin{array}{ll} \text{(sentence)} \rightarrow \text{(subject)}(\text{verb})(\text{complement}) & \text{(article)} \rightarrow \text{the} \\ \text{(subject)} \rightarrow \text{(article)}(\text{adjective})(\text{noun}) & \text{(adjective)} \rightarrow \text{little} \end{array}$$

The reader can easily construct a derivation of simple sentences like "The little boy went to school", etc.

In the study of formal languages, relation between the concept of recursively enumerable languages (or their equivalents in terms of Turing machines and / or normal algorithms) and formal grammars are of importance due to the following characterization of languages generated by formal grammars. Chomsky and Davis have shown that (see [4,6]):

- A language $L \subseteq A^*$ over the finite alphabet A is generated by a formal grammar $\Gamma = (V, A, \pi)$ if and only if L is recursively enumerable.

It is clear from the above that the languages generated by formal grammars are far too general to be of much practical importance. It is therefore customary to consider various special types of grammars and languages generated by them. The following is a classification of formal grammars $\Gamma = (V, A, \pi)$ in terms of their productions in π :

Type 0: This is the class of all formal grammars of the form $\Gamma = (V, A, \pi)$ defined above.

Type 1: This type of grammars are also called *context-sensitive*. They are characterized by the fact that every production in π has the form

$$u\alpha v \rightarrow uxv \quad (t1)$$

where $\alpha \in B$, $u, v \in B^*$ and $x \in V^+$.

Type 2: The grammar Γ is said to be *context-free* (or Type 2), if every production in π has the form

$$\alpha \rightarrow x \quad (t2)$$

where $\alpha \in B$ and $x \in V^*$. A language L generated by a context-free grammar is said to be *algebraic*.

Type 3: A grammar Γ is said to be *right linear* or *rational* (or Type 3) if every production in π has the form

$$\alpha \rightarrow x\beta \quad \text{or} \quad \alpha \rightarrow y \quad (t3)$$

where $\alpha, \beta \in B$ and $x, y \in A^*$. A language generated by a rational grammar is called a *rational language*.

We denote by \mathcal{L}_i , $i = 0, 1, 2, 3$ the class of all languages generated by grammars of Type i ; if $L \in \mathcal{L}_i$, we shall say that L is a Type i language. Thus, by Chomsky's result quoted above Type 0-languages are precisely recursively enumerable languages. It can be shown that Type 1 languages are recursive languages. Since recursive languages are recursively enumerable, it follows that $\mathcal{L}_1 \subseteq \mathcal{L}_0$. Since there exists recursively enumerable sets that are not recursive, there exists $L \in \mathcal{L}_0$ such that $L \notin \mathcal{L}_1$. Hence the inclusion $\mathcal{L}_1 \subseteq \mathcal{L}_0$ is strict so that $\mathcal{L}_1 \subset \mathcal{L}_0$. It is clear from the Equations (t1) and (t2) that any Type 2 language is Type 1. It can be shown that Type 2 languages are precisely those that are accepted by a *push-down automaton* and that there are Type 1 languages not accepted by any push-down automaton. Hence the inclusion $\mathcal{L}_2 \subseteq \mathcal{L}_1$ is strict. Similarly Type 3 languages are those accepted by a *finite-state automaton* and that not every push-down automaton is finite state. Hence we have the inclusion $\mathcal{L}_3 \subseteq \mathcal{L}_2$ which is also strict. Thus we have the following hierarchy of formal languages

$$\mathcal{L}_3 \subset \mathcal{L}_2 \subset \mathcal{L}_1 \subset \mathcal{L}_0$$

where the inclusions are all strict. This hierarchy of languages is called the *Chomsky hierarchy*.

Most of the existing studies are concentrated on the families of algebraic (context-free) and rational languages since many *decision-problems* have algorithmic solutions for these classes of languages. Thus

for example, in an algebraic language L , the problem $w \in L$ has a solution. We refer the reader to [1,3,8,9,12] for details of many concepts that were informally introduced above.

REFERENCES

- [1] M. A. Arbib (Ed.), *Algebraic Theory of Machines, Languages, and Semigroups*, Academic Press, New York, London.
- [2] E.R. Berlekamp, 1968, *Information Theory and Reliable Communications*, John Wiley & Sons, New York.
- [3] N. Chomsky, 1957, *Syntactic Structures*, Mouton, The Hague.
- [4] N. Chomsky, 1959, On certain formal properties of grammars, *Inf. Control.* **2** 137-167.
- [5] N. Chomsky and M. P. Schützenberger, 1963, Algebraic theory of contest-free languages in P. Braffort and D. Hirschberg (Eds.), *Computer Programming and Formal Systems* 118-161, North Holland Publishing Co., Amsterdam.
- [6] M. Davis, 1958, *Computability and Unsolvability*, McGraw-Hill Book Co., New York.
- [7] R. G. Gallager, 1968, *Algebraic Coding Theory*, McGraw-Hill Book Co., New York.
- [8] S. Ginsburg, 1966, *Mathematical Theory of Contest Free Languages*, McGraw-Hill Book Co., New York.
- [9] M. Gross and A. Lentin, 1967, *Notions sur les Grammaires Formelles*, Gauthiers-Villars. Paris.
- [10] P. R. Halmos, 1960, *Naive Set Theory*, D. Van Nostrand Co., Inc., Princeton, New Jersey.
- [11] S. C. Kleene, 1936, General recursive functions and natural numbers, *Math. Ann.* **112** 727-742.
- [12] G. Lallement, 1979, *Semigroups and Combinatorial Applications*, John Wiley & Sons, New York.
- [13] A. I. Mal'cev, 1965, *Algorithms and Recursive Functions*, Moscow; transl. Wolters-Noordhoff, 1970.
- [14] A. A. Markov, 1947, Impossibility of certain algorithms in the theory of associative systems, *Dokl. Akad. Nauk* **55** 587-590 (Russian).
- [15] A. A. Markov, 1947, Impossibility of certain algorithms in the theory of associative systems II, *Dokl. Akad. Nauk* **58** 353-356 (Russian).
- [16] A. A. Markov, 1951, Impossibility of certain algorithms in the theory of associative systems, *Dokl. Akad. Nauk* **77** 19-20, 953-956 (Russian).
- [17] A. A. Markov, 1956, *Theory of Algorithms*, Tr. Mat. Inst. Akad. Nauk SSSR, **42**, (Russian).
- [18] E. L. Post, 1936, Finite combinatory processes. Formulation, *J. Symb. Logic* **1** 103-105.
- [19] E. L. Post, 1947, Recursive unsolvability of a problem of Thue, *J. Symb. Logic* **12** 1-11.
- [20] F. de Saussure, 1931, *Course de Linguistique Générale*, 3rd ed., Gauthiers-Villars, Paris.
- [21] P. C. Suppes, 1960, *Axiomatic Set Theory*, D. Van Nostrand Co., Inc., Princeton, New Jersey.

- [22] A. Thue, 1914, Probleme über Veränderungen von Zeichenreihen nach gegebenen Regeln, *Skr. Vid. Krisitiania, I Mat. Naturv. Klasse* No.10 34pp.
- [23] A. M. Turing, 1937, On computable numbers with application to Entscheidungs problem, *Proc. Lond. Math. Soc.* **42** 230-265.

Department of Mathematics
University of Kerala
Kariavattom-695 581, INDIA

FORMAL LANGUAGES RELATED TO AUTOMATA

A. R. RAJAN

Department of Mathematics, University of Kerala

July 25, 1994

ABSTRACT: In this talk I give a brief idea of how languages can be specified by Automata and provide some simple properties of such languages. The special class of languages that is considered here is the class of Type three languages. Descriptions of this class in terms of rational expressions and grammars are also considered.

• Introduction

The study of formal languages has been progressing with two major objectives.

- (1) *to provide a mathematical setting for describing the structure of natural languages*
- (2) *To develop necessary tools for constructing special purpose languages such as programming languages*

Any natural language can be considered as the set of all correct sentences of the language and a theory or grammar for the language consists of providing rules for formation of all correct sentences. Abstracting these ideas we obtain formal languages and corresponding formal grammars (see [1], [3]).

• Preliminaries

Analogous to natural languages, a formal language can be described by specifying the basic vocabulary which we call the alphabet and acceptable strings of the alphabet (ie. sentences). The precise definitions are as follows (see [2],[3],[4], etc.)

Let V be any finite set. let us write $V = \{a, b, c, \dots\}$. Any finite string of elements of V (with possible repetitions) is called a word over V . The set of all words over V including the empty word (denoted by 1) is denoted by V^* . V^* is called the *Free Monoid* on V . It may be noted that

a monoid is an algebraic structure where a *multiplication* is defined on it. If w_1, w_2 are words over V (ie. $w_1, w_2 \in V^*$) then the product $w_1 w_2$ is defined as the word obtained by concatenation of w_1 followed by w_2 . Note that the empty word 1 will act as identity for this multiplication.

DEFINITION 1. Let V be any finite set. A subset L of V^* is called a language over V . We call V the alphabet for L .

EXAMPLE 1.

Let $V = \{cat, mouse, ate, the\}$ and L be the set of all *correct* sentences formed with the above words. Then with the conventional meaning of *correct* we see that

the cat ate the mouse

is in L whereas

the ate cat mouse

is not in L .

A language, if it is finite can be specified by enumerating all its words. But when the language is infinite one requires a generating device or a recognizing device to specify the language. One general method is to provide what is known as the generating grammar for the language.

DEFINITION 2. (cf. [4]) A generative grammar is an ordered quadruple $G = (V_N, V_T, X_0, F)$ where V_N, V_T are disjoint sets (called nonterminal alphabet and terminal alphabet respectively), $X_0 \in V_N$ and F a finite set of ordered pairs (P, Q) such that Q is a word over the alphabet $V = V_N \cup V_T$ and P is a word over V containing at least one letter (ie. member) of V_N .

Here X_0 is called the initial letter and elements $(P, Q) \in F$ are called productions. If $(P, Q) \in F$ then often we write $P \rightarrow Q$.

If $P = P_0, P_1, P_2, \dots, P_n = Q$ are words over V such that $(P_i, P_{i+1}) \in F$ for $i = 0, 1, \dots, n-1$ then we write $P \rightarrow^* Q$ and say that Q is derived from P by the given productions in G .

The language $L(G)$ generated by the above grammar is described as follows.

$$L(G) = \{P \in (V_T)^* : X_0 \rightarrow^* P\}$$

EXAMPLE 2.

Let

$$G = (\{X\}, \{a, b\}, X, \{X \rightarrow 1, X \rightarrow aXb\})$$

where 1 is the empty word. Then

$$L(G) = \{a^i b^i : i = 0, 1, \dots\}$$

where $a^0 b^0$ is taken as 1.

EXAMPLE 3.

Let

$$G = (\{X\}, \{a, b\}, X, \{X \rightarrow 1, X \rightarrow aXb, X \rightarrow bXa, X \rightarrow XX\})$$

Then

$$L(G) = \{w : \text{number of 'a' in } w = \text{number of 'b' in } w\}$$

If we omit the production $X \rightarrow 1$ then the corresponding $L(G)$ is empty.

Grammars and the corresponding languages they generate are classified by imposing suitable conditions on the form of their productions. This leads to classification of Type-0, 1, 2, 3 -grammars. Among these Type 3 grammar and the corresponding languages are amenable to many algebraic treatments.

• Rational Languages

We describe here a class of languages in algebraic terms and later see that this class coincides with the class of Type 3 languages.

If L, L' are languages over an alphabet V , then we define

$$LL' = \{uv : u \in L, v \in L'\}$$

and

$$L^* = \{u_1 u_2 \dots u_n : n \geq 1, u_i \in L \forall i\} \cup \{1\}$$

Writing $L^0 = \{1\}$ we see that $L^* = \cup_{n \geq 0} L^n$ where $L^n = L.L^{n-1}$ for $n = 2, 3, \dots$

Rational languages are defined in terms of certain expressions called rational expressions.

DEFINITION 3. Let V be an alphabet. Rational expressions over the alphabet V are defined recursively as follows.

- (1) Φ , 1 and a where $a \in V$ are rational expressions
- (2) If E, E' are rational expressions, then $E + E'$, $E.E'$ and E^* are rational expressions.

DEFINITION 4. To each rational expression E , a language $L(E)$ is associated which is given by the following rules:

- (1) $L(\Phi) = \Phi$, the empty language
- (2) $L(1) = \{1\}$, $L(a) = \{a\}$
- (3) If E, E' are rational expressions, then $L(E + E') = L(E) \cup L(E')$, $L(E.E') = L(E).L(E')$, $L(E^*) = (L(E))^*$

Rational languages are those languages $L(E)$ determined by a rational expression E . For example $L = V^*aV^*$ with $a \in V$ is a rational language as V^*aV^* can be considered as a rational expression also.

It is clear from the definition that the class of rational languages is closed under union, product and $*$ -product. The fact that it is closed under intersection and complementation are proved using more finer structure of these languages.

Now we look at this class of languages through a different perspective. This process is very much mechanical and the device is termed a *Automaton* (see[2] or [4] for details).

• Automaton

An automaton (precisely a finite deterministic automaton) is a 5-tuple

$$A = (S, V_T, s_0, S_1, F)$$

where S, V_T are finite sets called set of states and set of terminal alphabets respectively, $s_0 \in S$, called the initial state, $S_1 \subset S$ called the set of final states and $F : S \times V_T \rightarrow S$ called the set of transitions of A .

If F maps (s, a) to s' then we write $s.a = s'$. The map F can readily be extended to a map of $S \times (V_T)^* \rightarrow S$ as follows. For $s \in S$, set

- (1) $s.1 = s$
- (2) and if $w = a_1a_2 \dots a_k \in (V_T)^*$ then $s.w = (sa_1).(a_2 \dots a_k)$

This inductively defines $s.w$ for all $w \in (V_T)^*$. If $w = a_1a_2 \dots a_k$ and if $s.w = s'$, we describe this by a diagram as follows:

$$s = s_0 \xrightarrow{a_1} s_1 \xrightarrow{a_2} s_2 \rightarrow \dots \xrightarrow{a_k} s_k = s'$$

We call w a path in \mathbf{A} from s to s' .

We call a path in an automaton \mathbf{A} to be a successful path if it starts from the initial state s_0 and ends in one of the terminal states (ie. in an element of S_1). The set of all successful paths is called the language recognized by the automaton and is denoted by $L(\mathbf{A})$.

Several modifications on the definition of the automaton can be made without causing any change in the language it recognizes. One such modified version is called the *Non Deterministic Automaton*. Here instead of taking F as a mapping of $S \times V_T \rightarrow S$, F is taken as a subset of $S \times V_T \times S$.

• Consequences

When describing properties of rational languages, the question whether the complement of a rational language is rational was left unanswered. Now we see that using Automata this question can be easily answered.

RESULT 1. If L is recognizable by an automaton, then so is the complement of L .

Assuming that $L = L(\mathbf{A})$ where $\mathbf{A} = (S, V_T, s_0, S_1, F)$ consider the associated automaton $\mathbf{A}' = (S, V_T, s_0, S'_1, F)$ where $S'_1 = S - S_1$. Then clearly $L(\mathbf{A}') = L'$ where L' is the complement of L .

Now we describe an algorithm to determine whether a given recognizable language is nonempty, finite or infinite.

RESULT 2. (cf.[5]) Let \mathbf{A} be an automaton with n states. Then

- (1) $L(\mathbf{A})$ is nonempty if and only if there exists $w \in L(\mathbf{A})$ with length of $w < n$
- (2) $L(\mathbf{A})$ is infinite if and only if there exists $w \in L(\mathbf{A})$ of length k such that $n \leq k \leq 2n$.

• Algebraic Treatment

The languages described here can be treated in pure algebraic terms. We can associate a structured object "Monoid" or "Semigroup" with these languages in such a way that these monoids or semigroups characterize the languages.

DEFINITION 5. Let $L \subset V^*$ be a language. For $x \in V^*$, the set of all contexts of x in L is defined as

$$\text{cont}_L(x) = \{(u, v) \in V^* \times V^* : uxv \in L\}$$

The relation P_L defined on V^* by

$$xP_L y \text{ if and only if } \text{cont}_L(x) = \text{cont}_L(y)$$

is called the syntactic congruence of L .

In this case $M(L) = V^*/P_L$ (ie. the set of all congruence classes of P_L) is a monoid called the *syntactic monoid* of L . It is well known in algebra that the map $\eta_L : V^* \rightarrow V^*/P_L = M(L)$ defined by $w \rightarrow P_L(w)$ = the congruence class of w , is a multiplication preserving map (ie. homomorphism) from the monoid V^* into the syntactic monoid $M(L)$. The following theorem characterizes rational languages

THEOREM 1. (cf. [2]) Let $L \subset V^*$ Then L is rational if and only if $M(L)$ is finite.

It may be noted that in this case

$$L = (P)\eta_L^{-1}$$

where $P = (L)\eta_L$. Thus we arrive at the description of an algebraically recognizable language as follows.

DEFINITION 6. (cf. [2]) A language L over an alphabet V is said to be recognizable if there exists a finite monoid M , a homomorphism $\Phi : V^* \rightarrow M$ and a subset P of M such that $L = (P)\Phi^{-1}$.

There is a well known theorem (Kleene's Theorem) which states that all the different descriptions given above represents the same class of languages, ie. the class of rational languages or the class of Type 3 languages.

REFERENCES

- [1] N. Chomsky, 1957, *Syntactic Structures*, Mouton, The Hague.
- [2] M. Gross and A. Lentin, 1970, *Introduction to Formal Grammars*, Springer - Verlag.
- [3] G. Lallement, 1979, *Semigroups and Combinatorial Applications*, John Wiley & Sons, New York.
- [4] Saloma Arto, 1973, *Formal Languages*, Academic Press.
- [5] H. J. Shyr, 1991, *Free Monoids and Languages*, Lecture Notes, National Chung Hsing University, Taiwan.

JOSEPH'S CAT

T. MADHAVA MENON
Trivandrum

[The learned audience found the Endowment Lectures on Mathematical Linguistics, at the just-concluded All India Conference of Dravidian Linguists, somewhat strenuous. Later, a delegate said that the speakers started with Mathematics, and came to Linguistics as a special and limited application. This paper is an attempt to introduce certain elementary concepts of Formal Logic in a simple way, so that this requirement is to some extent satisfied. It stands in relation to the study of Formal Logic in the same way as familiarity with the English alphabet stands in relation to the study of Shakespeare. At a previous session, a delegate had suggested that it would be easier to generate sentences if an initial cue was given, eg., "Joseph's cat"; hence this title.]

The use of "Mathematical" concepts is now universal in learned discourse including Linguistics. The term "Mathematical concepts" includes the foundations of Mathematics and its meta-theory, Formal Logic and Set Theory. In this paper, we analyse the "simple" terms, "Joseph's cat", from a formal logical point of view.

The Propositional Calculus: Formal Logic starts, as does most mathematical discourse, with *axioms*, viz., unproven assertions involving undefined "primitive" concepts. The axioms lead to theorems, which can be proven within the framework of the axioms. There are only two conditions of validity or "values", viz., either "True" or "False". Multi-valued logic systems have been constructed with a gradation of values in between, and serve as the foundations of the theory of probability, but for the sake of simplicity, these systems are not dealt with here.

The "primitives" of the system are "propositions" formalized by the symbols "p", "q", "r", etc., consisting of lower case letters of the English alphabet, and "operations" among them, consisting of "." for conjunction (roughly equivalent to "and"), "v" for alternation ("or"), " \rightarrow " for implication ("if ..., then ..."), and " \sim " for negation ("not-p"). Thus, if p and q are propositions, then p.q means "p and q", $p \vee q$ means "p or q", $p \rightarrow q$ means "if p, then q" and $\sim p$, "not-p". Different authors have adopted several variations of these symbols; eg., an inverted v (" \wedge ") for conjunction, a double arrow (" \leftrightarrow ") for implication, a left-open bracket (" \supset ") for implication, and a forked line (" \neg ") for negation.

While most of these operations correspond to the significance they have in natural languages, the implication $p \rightarrow q$ is different in that it is

falsified only if the value of the "consequent", viz., q , is false. A statement is "meaningless" and therefore not a "proposition" if its value is not determinable as either True or False; it is a "Tautology" if a proposition is always true, eg., $p \rightarrow p$, and a "contradiction" if it is always false, eg., $p \rightarrow \sim p$. It has been demonstrated that these operations can be defined in terms of one another, or may be replaced by a single one, called "the Sheffer's stroke", symbolised by " $|$ ", otherwise known as "NAND" for "not-and". This finds considerable application in the design of computers and is a powerful tool in Boolean algebra; however, for the sake of simplicity, the inter-relationship among these operations is not discussed here.

A Truth Value table can be constructed as follows, where "T" stands for True, and "F" for False:

p	q	$p.q$	$p \vee q$	$p \rightarrow q$	$\sim p$	$\sim q$
T	T	T	T	T	F	F
F	T	F	T	T	T	F
T	F	F	T	F	F	T
F	F	F	F	T	T	T

It will be noticed that if both the first term (antecedent) and the second term (consequent) of an implication are False, then the implication is True! This is logically evident from the fact that the truth values are consistent with this analysis. A closer approximation to the natural language is shown in the operation called "strict implication", eg., p if and only if (also written as 'iff') q , symbolised by $p \leftrightarrow q$, or $p \equiv q$ which to be True requires that both p and q should be True.

From these primitive symbols, certain formulas can be defined as well-formed, some of which are listed as axioms. Rules are derived for inferring formulas from one or more formulas taken as premises. A theorem within such a system is a formula capable of proof through a finite sequence of well-formed formulas (wff) each of which is an axiom or is inferred from a wff. An interpretation that satisfies the axioms of the system is called a "model". Thus any "Language" is an *interpretation of a logical system of abstract structures of construction of, and deductions from, "primitive" symbols, governed by specific rules of formation and inference.*

The use of Truth Value tables has been generalized as an algorithm or decision procedure which, when "mechanically" applied to any set of propositions, can determine their validity or otherwise. This is possible only when the propositions are used as entities requiring no other specification. But propositions tell us nothing about themselves or anything other than themselves! When a quality or "predicate" is required to be expressed into a proposition, we leave propositional calculus and enter into the calculus of Predicates.

"First Order" Predicate Calculus: In order to become a "sentence", a proposition requires a "quantifier", specifying the spread of statements in which it is true or false. Thus the statement "Joseph's cat" is not a proposition as nothing is asserted about it; but if the statement is "Joseph's cat exists", then it becomes a proposition which can be "quantified", existentially.

In natural language, "Joseph's cat" may have several meanings, two of which are exemplified. It may mean a species of cat characterised by some features which set them apart from other objects, in the same way as "Russel's viper" means, not a viper belonging to Russel (least of all Bertrand Russel!), but a species of vipers. Here it is a method of naming, and says in effect: "There exists a set of cats/vipers which are named as "Joseph's cats"/"Russel's vipers". Alternatively, "Joseph's cat" means: "There is at least one cat such that it belongs to Joseph". In either of these meanings, it does not extend to the class including "all" cats, but only to a few of them, and asserts that the entity exists. In symbols:

$$\exists x (C_{(x \equiv \text{"Joseph's cat"})}) \quad (..1)$$

which means, (where ' \equiv ' means "is identified and named as"), "There is at least one 'x' such that it is named "Joseph's cat"". The symbol " \exists " is called an "Existential quantifier". However,

$$\exists x (C_x . B_{xj}) \quad (..2)$$

means "There is at least one 'x' such that 'x' is a Cat, and 'x' "belongs to" j (Joseph).

The "qualities" expressed by the letters "C" and "B" above are roughly (emphasize *roughly*) equivalent to the verbs of natural language and indicate "predicates". In the case of "C" in (..1) and (..2), it means "is cat". The "subject" of this predicate is only one, viz., "x". Such predicates are called "monadic". The predicate "B" in (..2) however requires two "subjects", viz., x and j, and hence is called "dy-adic". Depending on the number of subjects, we can have "n-adic" predicates for any number. In the terminology of Set Theory, a predicate can be considered as a "function" mapping from one Set constituting its "domain", to another, constituting its "range".

In a monadic predicate, the domain and the range are the same. But in a dyadic predicate, the predicate indicates the mapping of the elements in its domain, written first underneath it, over its range, written after it. Thus the expression " B_{cj} " in (..2) indicates that it is a mapping of the domain of the set "c" = cats over the range of the set "j" = Joseph's belongings. It may also be understood as the Intersection of the Set of Cats

and the Set of Joseph's belongings. The symbols in (..2) assert at least one "x", such that it is *both* a cat *and* Joseph's belonging, exists. Nothing more is asserted in (..1) and (..2) than that "x" as designated and circumscribed by the quantities indicated in the predicate exists. The statement is asserted in respect of only the specific number, not exhausting all, covered by it, viz., it may be one, or more than one, but not all of the "x"-s which are "C"-s. The existential quantifier may be considered roughly (emphasize *roughly*) similar to the "Determiner 'the'" in "Transformational Grammar".

When a quality is to be asserted over every member of a class of objects, then the "Universal quantifier" is used. It is symbolized by an inverted capital A (\forall), and sometimes by inclusion in brackets, eg., (x). The latter notation is adopted here. Thus the statement "All cats are felines" can be represented by:

$$(x) (C_x \rightarrow F_x) \quad (..3)$$

signifying that for all values of x, if x is a Cat, then x is a Feline.

It is possible to define the two quantifiers alternatively in terms of each other; (..3) above can be expressed as:

$$\sim (\exists x)(C_x \cdot \sim F_x) \quad (..4)$$

meaning that there is not a single x such that, x is a cat and not-Feline. The universal quantifier "implies" the existential quantifier by what is called the "law of Instantiation", which may be intuitively seen by the fact that a universal statement cannot be true unless there is at least one instance where it is true;

$$\{(x)(C_x \supset F_x) \leftrightarrow \exists x (C_x \cdot F_x)\} \quad (..41)$$

in other words, all cats are felines if and only if at least one cat is feline.

Similarly, the existentially quantified statement in (..2) can be rewritten as:

$$\sim (x)(C_x \cdot \sim B_{cj}) \quad (..5)$$

paraphrasable as "It is not the case of all x-s that x is a cat and x does not belong to Joseph".

Dyadic Relations: Both in Mathematics and in the analysis of Kinship terms, the logic of dyadic predicate calculus comes in very useful. Thus:

A symmetrical relation is when:

$$(x)(y)(E_{xy} \rightarrow E_{yx}) \quad (..6)$$

where the E can be interpreted as "being equal to", or "being the sibling of"; eg., "x is equal to/is the sibling of y, if y is equal to/is the sibling of x".

It is *asymmetrical* when:

$$(x)(y)(B_{xy} \rightarrow \sim B_{yx}) \quad (..7)$$

where B can be interpreted as "being bigger than", or "being the brother of"; eg., "x is bigger than y (y is not bigger than x)/x is the brother of y (y is not the brother of x) - of course this is the famous case where y is the sister of x!

It is *transitive* when:

$$(x)(y)(z)[(B_{xy}.B_{yz}) \rightarrow (B_{xz})] \quad (..8)$$

where also B can be interpreted as "being bigger than", or "being the brother of"; eg., "if x is bigger than y and y is bigger than z, then x is bigger than z"/"if x is the brother of y and y is the brother of z, then x is the brother of z".

It is *intransitive* when:

$$(x)(y)(z)[(F_{xy}.F_{yz}) \rightarrow \sim F_{xz}] \quad (..9)$$

where F can be interpreted as "being the father of"; eg., "if x is the father of y and y is the father of z, then x is not the father of z".

It is *reflexive* when:

$$(x) I_{xx} \quad (..10)$$

where I can be interpreted as "being identical to"; eg., "everything is identical to itself".

A relation which is reflexive, symmetrical and transitive is called an *equivalence* relation. This is very useful in number theory.

It has been proved that, unlike in the case of propositional calculus, *no decision procedure is possible even in the case of First Order Predicate Calculus*.

Higher level Predicate Calculi: When the predicates themselves are quantified and become the "subjects" of higher order predicates asserting "qualities of qualities", then higher level of calculi are called into use. Thus for example,

$$(C)(C_x \leftrightarrow F_x) \quad (..11)$$

which may be interpreted as "Being a cat is being a feline for all occurrences of cats". Also please note that in this notation, I have, for purposes of "economy of expression", put the letter C in the universal quantifier (brackets) in **Bold face Capital** to indicate that it is used at a higher level. Still higher levels are required for statements such as

$$L_1: (x)[(S_x.V_x) \rightarrow T_x] \quad (..12)$$

which may be interpreted to mean: "In Language L_1 , for all x , if x is a sentence in L_1 , then if x is V = in the word order SOV, then x is T = grammatically correct".

Recursively, a "grammatically correct sentence in language L_1 " may be defined:

$$L_1: (T)(S)(S_x \rightarrow V_x) \quad (..13)$$

ie., "In Language L_1 , T , = grammatically correct sentence, is such that for all S , = a sentence, if x is a sentence, x is V , = in the word order SOV".

Elementary Set Theory: Such extensions of the predicate calculi lead to the abstract entities such as "classes", and these to Set Theory, which is the basic discipline of Pure Mathematics. When quantifications extend over predicates, individuals/objects to which a predicate applies become members of a "class" defined by the applicability of those predicates. Thus the statement: "All cats are felines" can be interpreted to mean that, if x is a member of the class of cats, then it is a member of the class of felines". This is known as the Principle of Abstraction, and may be stated thus: "Every monadic predicate has a class as extension".

The concept was still further extended to cover "classes of classes", mainly by Russel and Whitehead in their classic *Principia Mathematica* (1911-1912). They used this concept to define a number as the class of classes having that number as its numerical predicate. However, the extension of this theory of inclusion of classes (cf the "Axiom of Specification") led to the famous Russel's Paradox, relating to the class of those classes which are not members of themselves. This led him to propose the "Theory of Types", according to which there is a hierarchy of classes, such that class inclusion implies a "higher" type.

Historically, Cantor postulated the concept of Sets such that (1) a Set is a grouping into a single entity of objects of any kind; and (2) given an object x , and a set A , x is either a member of A or is not. This concept was used for the explanation of cardinal numbers, but led to Cantor's Paradox, where the number of the subsets of a set is greater than the number of elements in the set. Later, a set of axioms known as the Zermelo-Fraenkel

axioms were propounded, in which two axioms were "axiom schemata", generating an indefinite number of axioms. The axioms were simplified by the von Neuman-Bernays-Gödel axioms, an interesting implication of which was that a "class" is defined as a "Set" if it is a member of some class, thus avoiding Russel's Paradox. However, the strict logicity and completeness of any formulation of a system of statements is proved to be impossible because of Gödel's famous theorems of Incompleteness, under which, if a system of statements S is consistent, then the consistency of S cannot be proved within S, and S is therefore incomplete. (The symbolism of Set Theory is now familiar at the High School level).

Application to "Language": Set Theory is a powerful system which can be used to study "almost all" categories which can be subjected to quantificational schema. "Language" as a means of communication can be quantified in various ways. When "strictly" quantified in the framework of a two-valued logic, a "Language" becomes "Formal". Any number of such formal languages can be constructed; when a language has to be subjected to "automatic" processes by a machine, then, the rigidity of the quantificational framework should be sufficient to avoid ambiguities, and be always consistent with its syntax or formal structure. "Natural" languages are at varying "statistical" distances away from such specifications, and present difficulties in machine processing, which can be reduced by structuring "hierarchies" of formal languages, each admitting greater degrees of "freedom" by building into the structure greater varieties of formation rules at each level in the hierarchy. "Philosophically" therefore, "Language" is a *human* achievement, and subserves the purpose intended for it by the humans concerned.

"Mathematical" Linguistics: The statistical study of some aspect of language use over a designated population is sometimes known as "Mathematical" linguistics; this has to be distinguished from the "Mathematical" study of the Set Theoretic formulations of the structure of the language itself.

THE DRAVIDIAN AND AFRICAN LANGUAGES

CLYDE A. WINTERS
Chicago, Illinois

Dravidian languages are predominately spoken in southern India and Sri Lanka. There are around 125 million Dravidian speakers. These languages are genetically related to African languages. The Dravidians are remnants of the ancient Black population who occupied most of ancient Asia and Europe.

In the sub-continent of India, there were several main groups. The earliest inhabitants of India were the Negritos, and this was followed by the Proto-Australoid, the Mongoloid and the so-called *Mediterranean* type which represent the ancient Egyptians and Kushites¹.

The Proto-Australoid race, Mongoloid race and Africoid/ Mediterranean skeletal remains were all found at Harappan sites. The speech of this group of Australoids is believed to be Austric, a specimen of this language survives in the Munda speech. (Thapar 1972, p. 26). The Africoid/Mediterranean group is associated with Dravidian culture.

The Negritos found the earliest culture in the Indus Valley at Mehrgarh in 6000 B.C. They had domesticated goats and sheep and grew cereals.

But India was never extensively urbanized before 2400 B.C. The antecedents of the Harappan culture are village sites of the Baluchistan hills, the Nal culture and the Makran coast to the west of the Indus delta.

The ancient inhabitants of the Indus Valley were the Nubian Dravidian speakers. Today, Brahui, Malto and Kurukh are the Dravidian languages spoken in the Indus area.

B.B. Lal (1963) the Indian Egyptologist has shown conclusively that the Dravidians originated in the Saharan area 5,000 years ago. He claims that they came from Kush, in the Fertile African Crescent and were related to the C-Group people who found the Karma dynasty in the 3rd millennium B.C. The Dravidians used a common black-and red pottery, which spread from Nubia, through modern Ethiopia, Arabia, Iran into India as a result of the Proto-Saharan dispersal. The Nubian Dravidians

1 Clyde A. Winters, "The Proto-Culture of the Dravidians, Manding and Summerians", *Tamil Civilizations* 3, No. 1, 1985, pp. 1-9.

were one of the cattle herding groups that made up the C-Group culture of Nubia Kush². Thundy has observed that:

"If Kashi the holy place (alias Varanasi) is a Dravidian word, the name was bestowed on the place by the Dravidians after their homeland of Nubia, Upper Egypt, which is called Kush and Kasi in the Semitic languages"³.

He added that:

"...I view the common myth of Isis and Osiris in Egypt and Kannaki and Kovalan in India only as supporting the widely held theory that the Dravidians of India came from the Mediterranean region, particularly from Nubia, for ethnically and linguistically, the peoples of the two regions are strikingly similar"⁴.

There is physical evidence which suggest an African origin for the Dravidians. The Dravidians live in South India. The Dravidian ethnic group includes the Tamil, Kurukh, Malayalam, Kannada (Kanarese), Tulu, Telugu etc.

The ancient Indo-Aryan writings make it clear that the Indians were dark-skinned (*varna*) and had flat noses (Durant 1935, p. 396). This fact is supported by the *Adi Tirāviḍar* (Old Dravidians) who are black as their African brothers with a difference in hair texture. In ancient Tamil poems they are described as *mamai* (black). In addition, the ancient Dravidians practised a matriarchal system in Kerala and South Kanara.

Aravānan (1980) has written extensively on the African and Dravidian relations. He has illustrated that the Africans and Dravidians share many physical similarities including the dolichocephalic indexes (*Aravānan* 1980, pp. 62-63), platyrrhine nasal index (*Aravānan* 1980, pp. 25-27), stature (31-32) and blood type (*Aravānan* 1980, pp. 34-35). *Aravānan* (1980, p. 40) also presented much evidences, for analogous African and Dravidian cultural features including the chipping of incisor teeth and the use of the lost wax process to make bronze works of arts (*Aravānan* 1980, p. 41).

2 K.P. Aravanan, "Physical and Cultural Similarities between Dravidian and African", *Journal of Tamil Studies*, No. 10, 1976, pp. 23-27:24.

3 Z. Thundy, "The Egyptian Osiris-Isis Myths and the Dravidian Cilappadikaram", *Tamil Civilization* 1, No. 2, 1983, pp. 83-90.

4 Ibid., p. 88.

CULTURAL UNITY OF THE DRAVIDIANS AND AFRICANS

The Dravidians have maintained their ancient African heritage. There are numerous affinities between Dravidian and Black African culture and languages⁵.

As in Africa, the Dravidians built both small and large vessels from a single log or planks tied together. This method of boat construction has been common in Africa since the rise of ancient Egypt, and continues today in East Africa, Chad and along the Niger river.

In both Africa and Dravidic India the people were organized into various *castes* or corporations. Many of the corporations such as that of the blacksmiths in Africa and India have corresponding names e.g., Wolof *Kamara* and Telugu *Kamara*.

There are similarities in agricultural techniques in Africa and India. For example, both groups used the hoe for tilling the ground, manuring the ground to fertilize crops, terracing irrigation and canal building. There are also affinities in animal husbandry, and even the names of animals. For example, sheep: Wolof *xar*, Brahui (Dravidian) *xar* ram; and cow: Wolof *nag*, Serere *nak*, Tamil *nāku* 'a female buffalo' and Tulu *naku* heifer.

There are also similarities between the Dravidian and African religions. For example, both groups held a common interest in the cult of the Serpent and believed in a Supreme God, who lived in a place of peace and tranquility⁶. There are also affinities between the names of many gods including Amun/Amma and Murugan⁷. Murugan, the Dravidian god of the mountains parallels a common god in East Africa worshipped by 25 ethnic groups called Murungu, the god who resides in the mountains⁸.

In addition, among the *Ati Tirāviḍar*, the system of inheritance passes from the uncle to his nephews, instead of to his sons (*maru makkal*

5 Clyde A. Winters, "The Genetic Unity of Dravidian and African languages and culture", *Proc. of the First International Symposium on Asian Studies*, 1979, Asian Research Service: Hong Kong, 1979, pp. 1105-1120.

6 Thundy, p.87; J.T. Cornelius, "Are Dravidians Dynastic Egyptians", *Trans. of the Archaeological Society of South India 1951-1957*, pp. 90-117; U.P. Upadhyaya, "Dravidian and Negro-African", *International Journal of Dravidian Linguistics* 5, No. 1, 1976, pp. 32-64:39

7 Upadhyaya, p. 40.

8 *Ibid.*, p.39.

tāyam) as in Africa⁹. And in both South India and the Western Sudan of Africa, the dead were buried in terracotta jars¹⁰.

LINGUISTIC EVIDENCE FOR INDO-AFRICAN LINGUISTIC UNITY

1.1. Many scholars have recognized the linguistic unity of Black African (BA) and Dravidian (Dr.) languages¹¹. These affinities are found not only in the modern African languages but also that of ancient Egypt. These scholars have made it clear that lexical, morphological and phonetic unity exist between African languages in West and North Africa as well as the Bantu group.

1.2. *K.P. Aravānan* (1976) has noted that there are ten common elements shared by BA languages and the Dr. group. They are (1) simple set of five basic vowels with short-long contrast; (2) vowel harmony; (3) absence of initial clusters of consonants; (4) abundance of geminated consonants; (5) distinction of inclusive and exclusive pronouns in first person plural; (6) absence of degrees of comparison for adjectives and adverbs as distinct morphological categories; (7) consonant alternation on nominal increments noticed by different classes; (8) distinction of completed action among verbal paradigms as against specific tense distinction; (9) two separate sets of paradigms for declarative and negative forms of verbs; and (10) use of reduplication for emphasis.

1.3. There has been a long development in the recognition of the linguistic unity of African and Dravidian languages. The first scholar to document this fact was the French linguist L. Homburger (1950, 1951, 1957, 1964). Homburger who is best known for her research in African languages was convinced that the Dravidian languages explained the morphology of the Senegalese group, particularly the Serere, Fulani group. She was also convinced that the kinship existed between Kannada and the Bantu languages, and Telugu and the Mande group. She is credited with the discovery for the first time phonetic, morphological and lexical parallels between Bantu and Dravidian.

For example, she noted that the Bantu infinitive end with a final *-a*, the subjunctive in *-e*, the preterit in *-i* or *-idi*, and the doer's name in *-i*, are all found with identical values in Kannada and other Dravidian languages. Homburger also found that in both the Bantu and Kannada languages, the causal suffix is *-is*.

9 U.P. Upadhyaya and S.P. Upadhyaya, "Les liens entre Kerala et l' Afrique noire tels qu' ils ressortent des survivances culturelles et linguistiques", *Bull. de l' IFAN*, T. 41, ser. B, No. 1, 1979, pp. 100-132:108.

10 Upadhyaya, p.40.

11 Winters, *Op cit*, pp. 1115-1120.

1.4. Tuttle (1932) also contributed to the investigation of links between African and Dravidian languages. In a short paper he wrote in the 1930s, he presents numerous lexical and grammatical parallels in Dravidian and Nubian.

1.5. One of the most interesting studies done to date on the links between African and Dravidian languages was the work of N. Lahovary (1963). Lahovary in his review of the possible link between the languages spoken by the founders of the major ancient civilizations, gives a stimulating discussion of cognates among various African languages and Dravidian (Dr.). He gives numerous lexical examples for the ancient kinship of the Dravidian group and BA languages, including ancient Egyptian, Hausa, Bantu, Nubian and Somali, to name a few.

1.6. By the 1970s, numerous scholars had moved their investigation on the links between Dr. and BA languages into the Senegambia region. Such scholars as Cheikh T. N'Diaye (1972) a Senegalese linguist, and U.P. Upadhyaya (1973) of India, have proved conclusively, Homburger's theory of unity between the Dravidian and the Senegalese languages.

1.7. C.T. N'Diaye, who studied Tamil in India, has identified nearly 500 cognates of Dravidian and the Senegalese languages. Upadhyaya (1973), after field work in Senegal, discovered around 509 Dravidian and Senegambian words that show full or slight correspondence.

1.8. As a result of the linguistic evidence, Congolese linguist Th. Obenga suggested that there was an Indo-African group of related languages. To prove this point we will discuss the numerous examples of phonetic, morphological and lexical parallels between the Dravidian group: Tamil (Ta.), Malayalam (Mal.), Kannada/Kanarese (Ka.), Tulu (Tu.), Kui-Gondi, Telugu (Tel.) and Brahui; and Black African languages: Manding (Man.), Egyptian (E.), and Senegalese (Sn.).

COMMON INDO-AFRICAN TERMS

ENGLISH	DRAVIDIAN	SENEGALESE	MANDING
MOTHER	<i>amma</i>	<i>ama, mēn</i>	<i>ma</i>
FATHER	<i>appan, abba</i>	<i>ampa, bāba</i>	<i>ba</i>
PREGNANCY	<i>basaru</i>	<i>bīr</i>	<i>bara</i>
SKIN	<i>uri</i>	<i>neguru, guri</i>	<i>guru</i>
BLOOD	<i>nettaru</i>	<i>deret</i>	<i>dyeri</i>
KING	<i>mannan</i>	<i>mānsa, omād</i>	<i>mansa</i>
GRAND	<i>bīra</i>	<i>būr</i>	<i>ba</i>
SALIVA	<i>tuppāl</i>	<i>tūdde</i>	<i>tu</i>
CULTIVATE	<i>bey</i>	<i>mbey</i>	<i>be</i>
BOAT	<i>kulam</i>	<i>gāl</i>	<i>kulu</i>
FEATHER	<i>sōge</i>	<i>sīge</i>	<i>sī, sigi</i>
MOUNTAIN	<i>kunru</i>	<i>tūd</i>	<i>kuru</i>
ROCK	<i>kallu</i>	<i>xēr</i>	<i>kulu</i>
STREAM	<i>kolli</i>	<i>kal</i>	<i>koli</i>

PRONOMINAL PARALLELS IN BA AND DR. LANGUAGES

LANGUAGE	I P. SG.	II P.	III P.	I P. PL.	II P.	III P.
Dravidian	<i>an, nā, ne</i>	<i>i</i>	<i>a</i>	<i>an, anu</i>	<i>a, ar</i>	<i>aru</i>
Somali	<i>ani</i>	<i>adigu</i>	<i>isagu</i>	<i>innagu</i>	<i>annagu</i>	
Nubian	<i>anni</i>	<i>ir</i>	<i>tar</i>	<i>u</i>	<i>ur</i>	<i>tar</i>
Bantu	<i>ni</i>	<i>u</i>	<i>a</i>	<i>tu</i>	<i>m</i>	<i>wa</i>
Manding	<i>na, n</i>	<i>i</i>	<i>a, e</i>	<i>alu</i>		
Hausa	<i>na</i>	<i>ka, kin</i>	<i>ya</i>	<i>mun</i>	<i>kun</i>	<i>un</i>
Wolof	<i>mā</i>	<i>ya</i>	<i>na</i>	<i>ne, pu</i>	<i>ngen</i>	<i>na</i>
Egyptian	<i>ink</i>	<i>ntk, ntt</i>	<i>ntf</i>	<i>inn</i>	<i>nttn</i>	<i>ntsn</i>
Elamite	<i>u</i>			<i>un</i>	<i>nun</i>	<i>r, ir</i>

1.10 In African languages and Dravidian, there is a system of five basic vowels and three-fold distinction of lip-rounded and unrounded, and a two-fold distinction of duration (short, long).

VOWEL SYSTEM OF DRAVIDIAN BLACK AFRICAN

<i>i</i>		<i>u</i>		<i>ī</i>		<i>ū</i>
<i>e</i>		<i>o</i>		<i>ē</i>		<i>ō</i>
	<i>a</i>				<i>ā</i>	

1.11. There is also phonetic correspondence. Similarity exists between the dull sonorous consonants, the nasal and labial series. This is especially true in the pronunciation of the sonorous consonants, e.g., clay: Malinke *banko*, Bambara *bogo*, and Telugu *banko-mannu*; Telugu *varu*, *vallu*, *vandru*, *vandlu*, (they, them), equal Mande/Manding *alu* or *aralu*.

1.12. In Dravidian and Black African, the *-f*, and *-b* are derived from *-p*, and therefore they are also interchangeable. The letters *-r*, *-i*, *-d*, and *-c*, *-s*, *-z*, are also interchangeable.

DEMONSTRATIVE BASES

LANGUAGE	PROXIMATE	DISTANT	FINITE
Dravidian	<i>i</i>	<i>a</i>	<i>u</i>
Mande	<i>i</i>	<i>a</i>	<i>u</i>
Fulani	<i>o</i>	<i>a</i>	
Serere	<i>e</i>	<i>a</i>	
Wolof	<i>i</i>	<i>a</i>	<i>u</i>

2. Ancient Egyptian and Dravidian: There are numerous corresponding lexical items in Egyptian and Dravidian languages; below are a few:

Language:	Egyptian	Dravidian	Language:	Egyptian	Dravidian
abscess	<i>bnw.t</i>	<i>pun</i>	abyss	<i>kiki</i>	<i>kedu</i>
to go	<i>hp</i>	<i>po</i>	with	<i>hr</i>	<i>-nnu</i>
build	<i>qd</i>	<i>kattu</i>	chief	<i>neb</i>	<i>nab grand</i>
great/noble	<i>bw</i>	<i>bal</i>	young	<i>hrd</i>	<i>kura</i>
house	<i>l</i>	<i>il, ll</i>	speak	<i>mdw</i>	<i>matu</i>
small	<i>sr, srr</i>	<i>siru</i>	to be	<i>wnn</i>	<i>unn-</i>

3. Dravidian and Mande languages: The Dravidian and Mande group of languages are very close. In fact, in Dravidian *mande* means "people". It would appear that the speakers of these languages lived in close proximity during the neolithic in the Fezzan region of Libya (Winters 1985b). Winters has proved that the Dravidian and Mande languages are genetically related and that speakers of these languages jointly colonized parts of Africa, Asia Minor and the Far East.

3.1. The Manding group of languages and Dravidian are very close. They share many grammatical and lexical similarities. In both these languages *-ka*, is used to represent 'to be', as well as a subjunctive. For example, in the Mande languages *ka* is a particle of different values, which corresponds to *kā*, the infinitive in Telugu of the verb *ag-uta* 'to become' (Ta. *aga*), e.g., Man. *a ka-nye* 'it is good', Tel. *ka valenu* 'it is necessary'. The same radical *ka*, represents the optative form in Telugu, e.g., *āpani mundara kani* 'how is labour given first place'; and in Manding a *k' a a barka d' i ma* 'it is god who gives blessings'.

3.2. In Dravidian, the suffixes *-ke*, *-ge*, *-ka* are used as the primitive verb 'to be' or 'to do'. They are usually used with abstract nouns, e.g., *ol* 'to reign', *ol-ka* 'domination'; *ose* 'to be content', *ose-ge* 'delight'; *nammu* 'to believe', *nammu-ka* 'confidence'. The Dravidian *-ke* corresponds to the Mande verb *ke* 'to do', which is often used with the suffix *-la*, to form derived nouns. For example in Manding, *sene* 'cultivation', *sene-li ke-la* 'cultivator'; and *tobi* 'to cook', *tobi-li ke-la* 'the cook'.

3.3. In Telugu, the suffix *-tu*, is used as the present participle while in the Mande group there is *-to* fulfilling the same function e.g., Tel. *chestu* 'made', Man. *tege* 'to cut', *tege to* 'cutting'.

3.4. Moreover Telugu *kani* 'not to be', corresponds to Mande *kana* the prohibitive negative participle and subjunctive, e.g. Man *a kana bugo* 'do not hit him'.

3.5. The past participle suffix in Tamil is *tu*, *-du* or *-i*, in Telugu we have *-i* and *-ti*. The *-tu* or *-ti* suffix of the Dravidian languages, corresponds to the Mande *-ti* or *-te* suffix used to form the negative sense, e.g., Man, *a ya* 'he is here', *a ti ya* 'he is not here'; *a be ta* 'he is coming', *a te ta* 'he is not coming'.

3.6. In the Dravidian languages, *-lu*, *-ru*, *-u*, are used for the construction of the plural. We thus see analogy in the formation of the plural *tense* in Dravidian and Mande e.g., Tel. *magadu* 'husband', 'man', *magalu* 'men'; Man. *mogo* 'husband', *mogolu* 'husbands'.

3.7. Analogy exists between Manding and Dravidian terms.

I. Consonantal Correspondence

English	Tamil	Manding
	s = / = s	
woman	<i>asa</i>	<i>musa</i>
	t = / = t	
fire	<i>ti</i>	<i>ta</i>
	I = / = I	
house	<i>lon</i>	<i>lu</i> 'family habitation'
	d = / = t	
law	<i>di</i>	<i>tili</i>
camp	<i>dagha</i>	<i>otagh</i>
forest	<i>kādu</i>	<i>tū</i>
	m = / = m	
mother	<i>amma</i>	<i>ma</i>
land	<i>man</i>	<i>ma</i> 'surface, area'
	k = / = k	
kill	<i>kol</i>	<i>ki</i>
man	<i>uku</i>	<i>moko</i>
	b = / = p	
great	<i>pal</i>	<i>ba</i>
	x = / = s	
sheep	<i>xar</i> 'ram'	<i>sara</i>
	c = / = s	
penis	<i>col</i>	<i>sol-ma</i>
abundant	<i>cal, sal</i>	<i>s'ya</i>

II. Full Correspondence

English	Dravidian	Manding
life	<i>zi</i>	'abundance'
clay	<i>banko-mannu</i>	<i>banko</i>
blacksmith	<i>inumu</i>	<i>numu</i>
lie	<i>kalla</i>	<i>kalon</i>
cultivation	<i>bey</i>	<i>be</i>
lord, chief	<i>gasa</i>	<i>kana, gana</i>
to recite	<i>sid, sed</i>	<i>siti</i>
great	<i>bal</i>	<i>ba</i>
to do	<i>cey</i>	<i>ke</i>
rock	<i>kal</i>	<i>kulu</i>
if, what	<i>eni</i>	<i>ni</i>
to cut	<i>teg</i>	<i>tege</i>

4.1. **Somali-Dravidian:** Many affinities exist between the Dravidian and the Somali languages. These affinities include similarity in phonetic systems, pronominal concordance, demonstrative and lexical items. In Dravidian,

the demonstratives are characterized by the opposition of the radical vowels *a*, *i* and *u*. The same type of opposition appears in Somali demonstrative postposition *-ka*, *-ki*, and *-ku*. The pronominal concordance is best typified in Dravidian and Somali (Som.) in the possessive pronoun I or me, e.g., Dr. *en* 'me'; Som. *-ani* 'me'. Also in Dravidian and Somali, they share similar terms to denote youth, e.g., Dr. *ar*, *ru* of the second person used to denote 'young'; and Som. *arur* 'youth'.

4.2. Analogy between numerous grammatical points occur within Dravidian and Somali. In Dravidian, to make the plural form, the suffix *-lu* is used. This corresponds to the Somali plural element *-o*, *yo*, which is joined to nouns to make the plural form e.g., *inan-ka* 'boy', 'son', *inammo* 'sons'; *maga'a* 'name', *maga yo* 'names'.

4.3. In Dravidian, the suffixes *-ke*, *-ge* and *-ka* are used as the primitive verb *to be*; it is employed with abstract nouns, e.g. *ose* 'to be content', *osa-ge* 'gladness'. This corresponds to the Somali suffix *-kar* 'can', 'be able', which is used with the infinitive, e.g., *wān sameyn kareūey* 'I was able to do (it)'.

4.4. The past tense in many Dravidian languages is formed by the suffix *-i*, *-ya* and *-ti*, especially in the case of Telugu. The use of the Dravidian *-ya* suffix corresponds to the Somali particle *-yey*, used to form the past tense. In Somali *-nayya* is used as the present continuous, e.g., *samēy* 'do' *wān* 'I': *wān samēy* 'I am doing, making', with the addition of the *yey* suffix we have *wān samayyey* 'I made' or *wān gēyyey* 'I brought'. This use of *-yey*, agrees with the Telugu use of both *-i* and *-ya*.

4.5. Somali *-ī*, is used to make the past tense of the definite article e.g., *ninku faraska bū dilayya* 'the man is beating the horse', *ninkī bā faraskī dilay* 'the man beat the horse'. As you can see from the example given above, the *-ī* element is joined to both the subject *ninkī* 'the man' and the object *faraskī* 'the horse'. This use of *ī*, is parallel to Dravidian *-i*.

4.6. In Somali, the continuous present is *ayya*, e.g., *wān furayya* 'I am opening'; and the present imperative for the plural is *-a*, e.g., *fur* 'to open', *fura* 'open'. The Somali *-a*, is also used with a noun to indicate the present even if the verb has the past tense, e.g., *kitāb* 'book', *safari* 'journey': *kitābkanu safarkaygī bū tilmān:ayya* 'this book describes the journey which I took'. This use of the Somali *-a*, and *ayya*, *to* is like the Tamil present termination elements *āl* and *ān*.

4.7. Also, in both Dravidian and Somali, the doers name is suffixed by the elements *-i* or *-ī*.

4.8. Dravidian and Somali cognates.

English	Dravidian	Somali
to take	<i>koḷ</i>	<i>gē, qād</i>
to become	<i>āgu</i>	<i>garow</i>
full	<i>ar</i>	<i>dereg</i>
to be	<i>ul</i>	<i>ol</i>
camp	<i>dagha</i>	<i>deg</i>
woman	<i>mag-wa</i>	<i>nāg-ta</i>
fire	<i>ta-gula</i>	<i>dab-ta</i>
with	<i>-nnu</i>	<i>na</i>
beg	<i>ira</i>	<i>bari</i>
tree	<i>cettu</i>	<i>ged</i>
lose	<i>ila</i>	<i>hallē</i>
yield	<i>kay</i>	<i>yēlo</i>
sickness	<i>allal</i>	<i>il-mo</i>
to suck	<i>nag-il</i>	<i>nug</i>
hunger	<i>gasi, kasi</i>	<i>gadzo</i>
female sex organ	<i>al-ku</i>	<i>a lol</i>
young man	<i>al</i>	<i>y'il</i>
ear	<i>kātu</i>	<i>deg-ta</i>
cheek	<i>katuppal, kadapu</i>	<i>lāb-tab</i>
cheek, jaw	<i>gandu, kanda</i>	<i>g-nd</i>
neck	<i>kural, gantalu</i>	<i>luqun-ta, hunguri-gaha</i>
heart	<i>karalu, kard</i>	<i>wadne-ka</i>
belly	<i>basaru, vayiru</i>	<i>alōl (sha)</i>
mother	<i>āy, ayya</i>	<i>hōyo-da</i>
father	<i>appa, appan</i>	<i>ābbe-ha</i>
sneeze	<i>cīntu</i>	<i>sīn, hindis</i>
village	<i>ūr</i>	<i>tūlo-da</i>
ciltivate	<i>bele, bey</i>	<i>bēr</i>
eat	<i>un</i>	
horse	<i>pari</i>	<i>faras-ka</i>

5.1. **Dravidian and Nubian:** There is visible correspondence between the Dravidian (Dr.) group and Nubian: Nile Nubian (NN), Kordofan Nubian (KN), Old Nubian (ON) and Modern Nubian (MN). The unity between these languages was first recognized by Tuttle and later elaborated on by Lohovary (1963).

5.2. There is a similarity between Dravidian and Nubian pronouns:

	I Sg.	I Pl.	II Sg.	II Pl.	III Sg.	III Pl.
Dravidian	<i>ān, na</i>	<i>nam</i>	<i>i, ni</i>		<i>tan</i>	<i>tam</i>
Nubian	<i>āni</i>		<i>ir</i>	<i>ur</i>	<i>tar</i>	<i>ter</i>

5.3. Dravidian and Nubia both use the suffix *-ku* or *-ko* for the diminutive. In these languages *-ke* is often used to denote smallness e.g., *ke* in Coptic, *kenna* in Dravidian, *kina* in Nubian.

5.4. The dative suffix in Dravidian *-ke*, *-ge*, *-ki*, *-gi* and *-ku* corresponds to the Nubian dative-accusative suffixes: ON *-ka*, MN *-gi* and KN *-gi*.

5.5. In both Tamil and Nubian, there are similar genitive endings: Ta. *-n* or *-in*, NN *n-*, *in*, and KN *en*, *nini*.

English	Dravidian	Nubian
Play	<i>ād</i>	<i>od</i>
sister	<i>akka</i>	<i>keg</i>
say	<i>an</i> , <i>en</i>	<i>onul</i>
woman, daughter	<i>asa</i>	<i>as</i>
elephant	<i>āne</i> , <i>enugu</i>	<i>onul</i>
bean	<i>avari</i>	<i>ogod</i>
water	<i>er</i>	<i>iri</i>
mother	<i>ia</i>	<i>ēn</i>
to be	<i>ir</i>	<i>in</i>
true	<i>olle</i>	<i>ale</i>
day	<i>ulla</i>	<i>ul</i>
son	<i>maga</i>	<i>ga</i>
mountain	<i>malē</i>	<i>mule</i>
fish	<i>mīn</i>	<i>anissi</i>
eat	<i>tin</i> , <i>ti</i>	<i>dī</i>
stone, rock	<i>kal</i>	<i>kulu</i>
tongue	<i>na</i>	<i>nar</i> , <i>nad</i>
shore	<i>kare</i>	<i>gār</i>

6.1. **Dravidian and Senegalese:** Cheik T. N'Diaye (1972) and U.P. Upadhyaya (1976) have firmly established the linguistic unity of the Dravidian and Senegalese languages. They present grammatical, morphological, phonetic and lexical parallels to prove their point.

6.2. In the Dravidian and Senegalese languages there is a tendency for the appearance of open syllables and the avoidance of non-identical consonant clusters. Accent is usually found on the initial syllable of a word in both these groups. Upadhyaya (1976) has recognized that there are many medial geminated consonants in Dravidian and Senegalese. Due to their preference for open syllables final consonants are rare in these languages.

6.3. There are numerous parallel participles and abstract noun suffixes in Dravidian and Senegalese. For example, the past participle in Fulani (F) *-o*, and *ōwo* the agent formative, corresponds to the Dravidian *-a*, *-aya*, e.g., F. *windudō* 'written', *windōwo* 'writer'.

6.4. The Wolof (W) *-āy* and Dyolo *ay*, abstract noun formative corresponds to Dravidian *ay*, W. *bāx* 'good', *bāxāy* 'goodness'; Dr. *apala* 'friend', *bapalay* 'friendship'; Dr. *hiri* 'big', *hirime* 'greatness', and *nal* 'good', *nanmay* 'goodness'.

6.5. There is also analogy in the Wolof abstract noun formative suffix *-it*, *-itt*, and Dravidian *ita*, *ta*, e.g., W. *dog* 'to cut', *dogit* 'sharpness'; Dr. *hari* 'to cut', *hanita* 'sharpness'.

6.6. The Dravidian and Senegalese languages use reduplication of the bases to emphasize or modify the sense of the word, e.g., D. *fan* 'more', *fanfan* 'very much'; Dr. *bēga* 'quick', *bēgabēga* 'very quick'.

6.7. Dravidian and Senegalese cognates.

English	Senegalese	Dravidian
body	<i>W. yaram</i>	<i>uru</i>
head	<i>D. fuko, xōx</i>	<i>kukk</i>
hair	<i>W. kawar</i>	<i>kavaram</i> shoot
eye	<i>D. kil</i>	<i>kan, khan</i>
mouth	<i>D. butum</i>	<i>bāyil, vāy</i>
lip	<i>W. tun, F. tondu</i>	<i>tuti</i>
heart	<i>W. xol, S. xōr</i>	<i>karalu</i>
pup	<i>W. kuti</i>	<i>kutti</i>
sheep	<i>W. xar</i>	<i>ram</i>
cow	<i>W. nag</i>	<i>naku</i>
bronze	<i>W. xanjar</i>	<i>xancara</i>
skin	<i>dol</i>	<i>tōl</i>
mother	<i>W. yāy</i>	<i>āyi</i>
child	<i>D. kunil</i>	<i>kuññu, kūci</i>
ghee	<i>o-new</i>	<i>ney</i>

We provided above linguistic examples from many different African Supersets (families) including the Mande and Niger-Congo groups to prove the relationship between Dravidian and Black African languages. The evidence is clear that the Dravidian and Black African languages should be classed in a family called Indo-African as suggested by Th.Obenga. This data further supports the archaeological evidence accumulated by B.B. Lal (1963) which proved that the Dravidians originated in the Fertile African Crescent.

Thundy believes that the Dravidians may have left Nubia after Senefru (c.2613 B.C.) conquered Nubia¹². Senefru's raid caused much destruction and may have encouraged many Kushites to flee Nubia for safe areas of settlement.

HARAPPAN

The Harappans spoke a Dravidian language. According to Walter Fairservis, the ancestors of the Harappans came from Iran and

12 Thundy, *op cit*, p. 89.

Baluchistan. Fairservis makes it clear that the early cultures of Baluchistan are analogous to Early Dynastic Sumerian and Elamite sites.

Both pre-Harappan and mature Harappan sites occur on the Makran coast and parts of Iran. Pre-Harappan is the term applied to Harappan sites found in Baluchistan and Makran.

Pre-Harappan sites are typified by villages of the Nal-Amri period, and its equivalents from Punjab to Sind. The style of art from these sites is called Kulli-Nal, Amri and Kot Dijian. It was on the Makran coast that we find the earliest distinctive black-and-red ware common to Harappan sites and the terracotta bulls and symbols common to the mature Harappan phase.

The civilizations of the Indus have been divided into three periods, the *Pre-Urban Harappan Phase* (3200-2600 B.C.), *Urban (Mature) Harappan Phase* (c.2600-1700 B.C.) and the *Post Harappan Phase*.

The Indus region is an area of uncertain rains because it is located in the fringes of the monsoon. Settlers in the Indus Valley had to suffer both frequent droughts and floods.

Archaeological and linguistic evidence indicates that the Dravidians were the founders of the Harappan culture which extended from the Indus Valley through northeastern Afghanistan, on into Turkestan. The Harappan civilization existed from 2600-1700 B.C. The Harappan civilization was twice the size of the Old Kingdom of Egypt. In addition to trade relations with Mesopotamia and Iran, the Harappan city states also had active trade relations with the Central Asian peoples. (Winters:1990)

Fairservis (1975) makes it clear that early cultures of Baluchistan are analogous to Early Dynastic Sumerian, this movement eastward of the ancient Kushites led to the rise of the Indus cultures.

The Sumerians probably called the Indus Valley *Dilmun*. Dilmun was a rich trade centre that provided Sumer with many valuable trade items.

Archaeologists believe that there was a punctuated urbanism in the Indus Valley. This is indicated by the "short, rapid, even explosive" rise of the mature Harappan phase between 2600 and 2500 B.C. (Possehl & Raval:1989). Possehl (1990, p. 262) wrote that "Rather than emerging from a long slow period of gradual and constant cultural growth and modification leading to a slowly emerging pattern of urbanization and social complexity, the Harappan Civilization seems to have resulted from a short period of transformation, a veritable revolution within the urbanization process".

The Mature Harappan civilization is divided into two variants, the Sorath Harappan and the Sindhi Harappan. The Sindhi Harappan sites are sites characterized by elaborate architecture, fired brick construction, sewage systems and stamp seals. The Sindhi Harappan styles are found in Gujarat, Kutch, Punjab, Haryana and Uttar Pradesh. The major Sindhi cities include Mohenjodaro, Lothal, Rangpur, Harappa, Desalpur, Shirkotada, Manda. Ropar, Kalibangan and Chanhudaro.

The Sindhi Harappans possessed writing, massive brick platforms, well-digging, a system of weights-and-measures, black-and-red ware (BRW), metal work and beads. (Possehl 1990, p. 268)

The BRW tradition originated in Nubia and spread first to the Indus Valley and thence to India. The earliest occurrence of BRW in South Asia, occurs on the Kathiawad peninsula, parallel ware has been found at the lowest levels of Harappa and Lothal dating to 2400 B.C. (Rao 1972). Nayar (1977) has shown that the Harappan BRW has affinities to predynastic Egyptian and West Asian BRW dating to the same period. Rao (1972) has established the unitary nature of the BRW industry from Nubia to Asia.

The Harappans were masters of hydraulic engineering. They were a riverine people that practised irrigation agriculture. They had both the *shaduf* and windmills.

At Harappan sites, domestic quarters and industrial areas were isolated from each other (Fairervis:1991).

Sorath Harappan sites lack stamp seals, ornaments and elaborate architecture. Sorath is the ancient name for Saurashtra. The Sorath Harappan sites are located in Saurashtra, Kulli and the Harappan styles of Baluchistan and Gujarat.

The Indus region is an area of uncertain rains because it is located on the fringes of the monsoon (Fairervis:1987, p. 47). Settlers in the Indus Valley had to suffer frequent droughts and floods. Severe droughts frequently occurred in the Indus Valley. So, the people dug wells to insure for themselves a safe supply of water.

To compensate for the adverse ecological conditions, the Harappans settled along the Indus river (Fairervis:1987, p. 48.).

The Dravido-Harappans occupied over 1,000 sites in the riverine Indus Valley environments where they had soil and water reserves. The Harappan sites are spread from the Indus Valley to Ai Kharnoum in northeastern Afghanistan and southward into India. In Baluchistan and Afghanistan, Dravidian languages are still spoken today. Other Harappan

sites have been found scattered in the regions adjacent to the Arabian sea, the Derajat Kashmir and the Doab.

The Harappans were organized into chiefdoms, averaging between two and five acres (Fairservis:1987). The Harappans were a sedentary-pastoral people organized into various corporations such as sailor-fishermen, smiths, merchants and farmers. The Harappans also possessed the social technology of writing seals.

The Harappan sites are small and occupy only a few acres with little depth. This suggests that the Dravidian speaking colonists settled the Indus Valley over a period of a few decades (Fairservis:1987, p. 46). Fairservis (1987, p. 47) has shown that the site of Mohenjodaro was occupied for around 200 years.

Many archaeologists are beginning to accept the fact that the Harappan civilization was founded elsewhere and taken to the Indus Valley by the Harappan people. (Possehl & Raval:1989; Possehl:1990)

The Sumerians called the Indus Valley as *Dilmun* or *Tilmun* according to Sumerologist S.N. Krammer in The Sumerians: Their History, Culture and Character. Other specialists have begun to popularize the idea that the Indus Valley was called Meluhha, because of the Aryan mention of Meluhhaites in India when they arrived. There were Meluhhaites in India living along the Ganges, but these Meluhhaites were settled in India after Sestrosis I of Egypt conquered the Ganges region. It was also around this time that the Egyptians established colonies in Colchis near the Black Sea.

During the times of Sargon the Great of Sumer, Dravido-Harappan ships from Dilmun were anchored at Agade docks in Mesopotamia. The ships of Dilmun exported gold, copper utensils, lapis lazuli, ivory, beads and semiprecious stones.

Today, there are isolated pockets of Dravidian-speaking groups surrounded by Indo-Aryan speakers. Dravidian languages are spoken by tribal groups in Gujarat, Maharashtra, Madhya Pradesh, Orissa, West Bengal and Bihar.

The Harappans were fine engineers and craftsmen. They built large cities with complex drain systems under the streets in some of their cities.

The Harappans cultivated wheat, barley and millet. They had domesticated sheep/goats and cattle.

Mainly sedentary-pastoralist, the Harappans had a highly developed grain storage system. The main cities of the Harappan civilization were Mohenjodaro, Harappa, Chanhudaro, Kalibangan and

Like Gundert's *Malayalam-English Dictionary*, Kittel's *Kannada-English Dictionary* is still, to this day, unsurpassed.

Finally I would like to mention Kittel's erudite study *On the Origin of the Linga Cult in India* (Über den Ursprung des Lingacultus in Indien), Basel Mission Mangalore 1876, which carries most valuable informations on ancient Hindu religious and metaphysical aspects.

Many other German scholars in the field of Dravidology could still be mentioned, thus, for instance, one of Ziegenbalg's successors at Tranquebar Johannes Ph.Fabricius with his *Malabar and English Dictionary* (in co-authorship with Breithaupt) of 1779; or Johann P. Rottler with his *Tamil-English Dictionary* (based on Fabricius) of 1834; and more recently Hermann Beythan, author of a grammar of modern spoken Tamil: *Praktische Grammatik der Tamilsprache*, 1943.

Eugen Hultzsch (1857-1927), from 1886 to 1903 Epigraphist to the Government of Madras, Examiner of Sanskrit and Fellow of the University of Madras, would certainly deserve more detailed mentioning, mainly as the editor of *South Indian Inscriptions* (3 vls., Madras 1890-1903), most of them collected from Tanjore and North Arcot Districts, i.e., the area of Mahabalipuram and Kancheepuram, and also with his *Reports on Sanskrit Manuscripts in South India* (3 vls., Madras 1895-1905).

Like him Otto Schrader (1876-1961), a master of all four Dravidian languages, Tamil, Malayalam, Kannada and Telugu, should evoke our special attention. Under his directorship from 1905 to 1914 (the outbreak of World War I), the Library of the Theosophical Society at Adyar/Madras gained its reputation as a scholarly institution of the first rank. But to deal with Hultzsch and Schrader would call for a separate lecture, which could perhaps be done on a later occasion.

When in 1953-54 Prof. V. Raghavan, the renowned Sanskritist of Madras University, had toured Indological Centres in Europe, specially in search of hitherto unknown or unclassified Sanskrit texts, he wrote a detailed report on his findings and experiences under the title *Sanskrit and Allied Indological Studies in Europe*, Madras 1956.

In the introductory note to the chapter on Germany we read:

Though the credit of discovering Sanskrit for the West goes to England and though it was at the feet of French pioneers that the first German Sanskritists like Bopp and Schlegel sat, Germany had taken to Sanskrit studies with such enthusiasm and disinterested love, that outside India, one can say without exaggeration, it has been a second home of

Sanskrit... To give an account of German contributions to Sanskrit would be to write the history of Sanskrit studies in modern times.

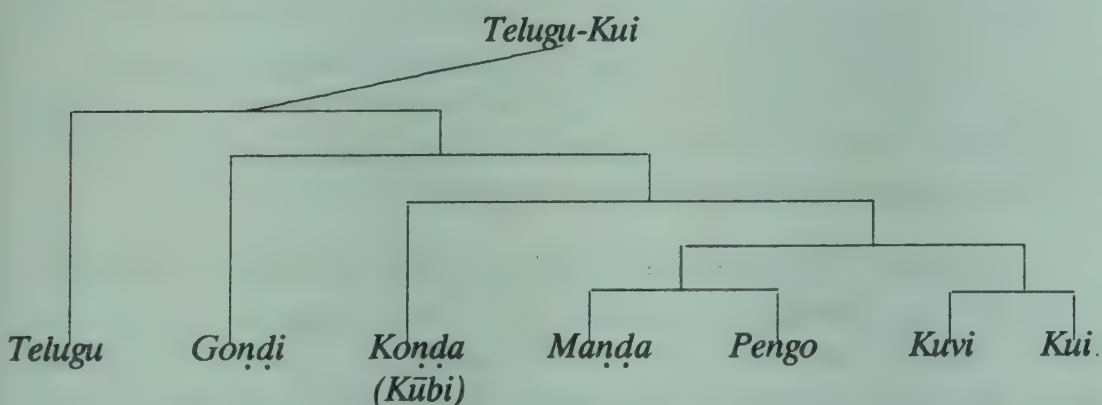
I think also that without exaggeration, similar statements could be made as regards Dravidian studies, which I wanted, if not to prove, at least to indicate in my lecture.

[Permission for reprint given by the Indo-German Society, Manipal, and by the author who first delivered it as a lecture to the said Society, Manipal on 21 August, 1993, is acknowledged.]

PLURAL MARKERS IN KUI AND KUVI: A RECONSIDERATION

PANCHANAN MOHANTY
Hyderabad University

1. The categories of gender and number are "inseparably connected in Dravidian" (Caldwell:1974:221). In other words, one cannot discuss either of them without taking the other into account. Kui and Kuvi are no exception to it. Kui and Kuvi together with *Kūbi* (or *Koṇḍa*) are three tribal languages belonging to the South Central branch of the Dravidian family (Zvelebil:1990:56). The names of these three languages consist of /ku/ 'hill' and the suffix /-i/ denoting 'man', and thus, each of them mean 'a hill man'. Previously, scholars were of the opinion that these "... three are very closely related languages." (Israel:1979: XXI) and "...the similarity in the three native names suggests their original identity as a single speech community, perhaps also culturally homogeneous." (Krishnamurti: 1969a:3). But according to Subrahmanyam (1969a:133) and Krishnamurti (1969b:327), Kui and Kuvi are sister languages whereas *Kūbi* should be a cousin of theirs. Zvelebil (1990:57) holds the view that *Kūbi* is a distant cousin of Kui and Kuvi. It is clear in his subgrouping of these languages which is as follows:



Thus, there is a difference of opinion regarding the relationship of *Kūbi* with Kui and Kuvi. An analysis of the plural markers of these three languages support Subrahmanyam and Zvelebil, not Israel and Krishnamurti.

Kui, Kuvi and *Kūbi* each has two numbers, i.e., singular and plural like other Dravidian languages. The former is generally unmarked whereas the latter, marked. That means the plural nouns are formed by adding

plural markers to the singular stems. Scholars like Andronov (1979), Krishnamurti (1975), Shanmugam (1972), Subrahmanyam (1969a, 1969b), and Zvelebil (1977, 1990) are unanimous on the point that Kui, Kuvi and *Kūbi* have two genders, i.e., masculine and feminine. With reference to *Kui*, Winfield (1928) follows the same line of argument in the beginning and divides all nouns into two categories like masculine and non-masculine. But immediately after that he tends to recognise "... a feminine gender as forming a separate section of the other-than-masculine class. For example, the simple nouns denoting female human beings have a somewhat different plural ending (-ska instead of -ka) from the rest of the non-masculine nouns." (Winfield 1928:12). Thus, he postulates three genders, i.e. masculine, feminine and neuter in his grammar of *Kui*; and other scholars like Dash (1981) and Chetty (1984) follow him.

While discussing the category of number in Kuvi, Israel (1979:47) recognises three genders in this language and states, "The plural nouns are mostly marked and the plurality is represented by three different groups of suffixes, of which the first one is exclusively for masculine nouns, the second one is mainly for the neuter nouns and the third one is exclusively for the feminine nouns. Thus plurality may be discussed under the above different genders, viz. masculine, neuter and feminine. There are two genders, masculine and non-masculine or neuter". But Krishnamurti (1969a:224) clearly states, "There are two genders, masculine and non-masculine or neuter." in *Kūbi*. Interestingly, when /-ska/ is recognised as the feminine plural suffix in both *Kui* and *Kuvi* by the scholars mentioned above and it is widely attested in these two languages, Krishnamurti (1969a: 226) cites only one instance, i.e., /koronali/ 'a nursing mother, a woman in birth pollution' which takes /-sku/ in the plural, i.e., /koronal-sku/ in *Kūbi*. This is a striking difference between *Kui* and *Kuvi*, on one hand, and *Kūbi*, on the other. For this reason, I have chosen *Kui* and *Kuvi* together to discuss their plural markers in this paper.

It should be noted here that Andronov (1977), Krishnamurti (1972), Subrahmanyam (1969a) and Zvelebil (1977, 1990) among others are of the opinion that *Kui* and *Kuvi* have only two genders, i.e., masculine (or high class) and non-masculine (or low class). But according to Chetty (1984), Dash (1981), Winfield (1928) and Israel (1979), there are three genders, i.e., masculine, feminine and neuter in these languages. Thus, the views of both the groups of scholars are contradictory. The compelling reason for the latter to recognise the feminine as a gender is obviously the feminine plural marker /-ska/ which they could not account for in any other way. So, the present paper aims to reexamine the existing data and analyses of *Kui* and *Kuvi* and give a more acceptable solution to the problem outlined above.

2. Let us now look at the plural markers of the languages concerned. There is an agreement among the scholars that there are only four plural morphemes in Kui and they differ from each other with respect to their individual shapes. According to Brierson (1963:460), they are /-ru/, -wi/, -ga/, -ska/. Winfield (1928:13-18) and Dash (1981:62) state that they are /-ru/, -ka/, -nga/, -ska/. In the data collected by me, the plural markers are the same as those given in Winfield (1928) and Dash (1981). Notice that Brierson's /-wi/ is not attested in anybody else's data. Again, even though /-ga/ is found, it is only a conditioned variant of /-ka/. That is, if there is a nasal in the stem-final position and /-ka/ is added to it, the [k] of /-ka/ changes to [g] by assimilating the feature [+voiced] of the preceding nasal. For example:

penu + *ga* → *penga* 'gods'
mīnu + *ga* → *mīnga* 'fishes'
Kanu + *ga* → *kanga* 'eyes'

It should be pointed out here that *Kui* retains the archaic Dravidian characteristic of not allowing a word to end in a consonant. Therefore, an enunciative vowel, viz. [u] is found at the end of certain nouns which end in consonants, and it gets deleted when the plural morpheme /-ka/ is added to these nouns to get the plural forms. Thus, the enunciative [u]-deletion and voicing of the [k] in /-ka/ are in feeding order.

The only other difference is that Chetty (1984) has /-ra/ whereas all others have /ru/. This difference is most probably dialectal, because, unlike others, Chetty has collected data from the district of Koraput. So, I will take only the plural markers /ru/, -ka/, -ska/, -nga/ of *Kui* for consideration.

The situation in *Kuvi* is quite similar. It is reported to have /-ri/, -ka/, -nga/, -na/, -n/, -yā, \tilde{V} (nasalization), -ska/ as plural markers (Israel: 1979:46-60). Out of these, /-ri/ and -ska are employed to pluralise only the masculine and the feminine stems respectively; and among others /-ka/, -nga/, -yā are quite frequently used as the neuter plural markers, though sometimes they are added to the masculine stems also. As one can see, /-na/, -n/, -yā, \tilde{V} (nasalization) / are variants of /nga/. /-na/ is the result of the deletion of the /g/ in /-nga/, and /-n/ is formed by deleting the /a/ in /-na/. /-yā/ is actually /-ā/ which is formed by the deletion of the /n/ in /-na/ and nasalization of the following /a/. As *Kuvi* does not allow 'vowel clusters' (Israel 1979:8), the /y/ glide is inserted between the final vowel of the stem and /ā/ and thus, it becomes /-yā/. This marker is an innovation in *Kuvi* as it is not attested in any other South Central Dravidian language. Finally, /- \tilde{V} / is created due to the deletion of the marker /-n/ which leaves its trace in the form of the nasalization of the preceding vowel in the stem. For these reasons, I will consider only /-ri/, -ka/, -nga/, -ska/ among the plural

markers of *Kuvi* in this paper. Notice that except the vowel /i/, in /-ri/, all these markers of *Kuvi* are identical with those of *Kui*.

3. First, let us look at the plural markers of *Kui*. According to Winfield (1928:13), "The plural of masculine words is formed by adding /ru/ to the singular form..." and he lists four exceptions in which either /-ka/ or /-nga/ is used in the place of /-ru/. Regarding the feminine nouns, he states that "The plural of feminine is formed by adding /-ska/ to the singular form." (Winfield:1928:14), and there is no exception to it. He divides the non-rational beings and inanimate objects into five classes according to the final vowel of the singular stems and formulates rules for the distribution of /-ka/ and /-nga/. But, all these rules except the two concerned with the stems ending in [o] and [e] have so many exceptions that one would hesitate to call them rules.

On the other hand, Dash (1981:72) states "The plural morpheme suffixed to nouns has four variants (allomorphs) /-ru ~ ka ~ ska ~ nga/.

The distribution is as follows:

1. Masculine plural /-ru ~ ka ~ ska ~ nga/
2. Non-masculine (+ human) /-ska/
3. Non-masculine (-human) /-ka ~ nga/."

Following Winfield (1928), he tries to determine the distribution of /-ka/ and /-nga/ on the basis of the stem-final vowels of singular nouns, and concludes, ".... the distribution of the plural allomorphs /-ka/ and /-nga/ is unpredictable." (Dash:1981:72).

Chetty (1984) also follows a similar line of analysis. Like Winfield (1928) and Dash (1981), he states that /-ra/ and /-ska/ are masculine and feminine plural markers respectively though some masculine nounstems are found taking the neuter plural markers /-ka and /-nga/ and describes the distribution of /-ka/ and /-nga/ taking the syllabic structures of the neuter noun stems, the vowels that occur at the end of those stems, and the composition of the ultimate syllables (Chetty:1984:89-105). But taking the exceptions into account, he concludes that "/-ka/ and /-nga/ are to be taken as morphologically conditioned (or lexically specific) allomorphs as they occur in the same phonological environment." (Chetty:1984:105).

Israel (1979) illustrates that *Kuvi*, like *Kui*, uses the plural markers /-ri/ and /-ska/ exclusively for the masculine and the feminine nouns; and the remaining ones like /-ka/ and /nga/ along with its variants are added to both the neuter as well as the masculine nouns.

What is significant here is that *Kui* and *Kuvi* are almost identical with respect to the use of the plural markers. All the above scholars agree that /-ru/ of *Kui* and /-ri/ of *Kuvi* are added only to the masculine stems and

/-ska/ is added only to the feminine stems; and /-ka/ and /-nga/ are used with both the masculine and the neuter stems.

The masculine markers /ru/ of Kui and /-ri/ of Kuvi can clearly be related to the Proto-Dravidian plural marker for human nouns *-Vr. According to Subrahmanyam (1969b:94), "The restriction of the -r Pl. suffix in the C(entral) Dr(avidian) nouns must be secondary since the alignment of the fem. with neu. in the Pl. category has to be interpreted as an innovation". Then, a comparison of all other plural markers, viz. /-ka, -ska, -nga/ clearly shows that all of them have /-ka/ in common. This, in fact, vindicates the position that originally there were only two plural markers in the proto-stage of the languages concerned, i.e., */- rV/ for masculine stems and */-ka/ for non-masculine stems, and "The separation between the feminine and the neuter genders seem to be later." (Zvelebil:1977:11). If /-ka/ is accepted as the retention of the proto-stage non-masculine plural marker, then we will have to account for the /-s-/ of /-ska/ and the /n-/ of /-nga/. Let us now discuss them one by one.

5. The /-sk/ plural marker is a characteristic of the South Central Dravidian languages (Krishnamurti:1972:258). In the related Central Dravidian languages like Kolami, Naiki, Ollari and Parji, the most common feminine plural marker is /cil ~ sil/ (Subrahmanyam:1968:177, 1969b:98; Tyler:1975:101). If we compare the so-called feminine plural markers of both these groups, we will notice that the common element between them is /-c-/ or /-s-/. While talking about the plural markers of the said languages, Subrahmanyam (1969b:99) states, "Even though it is not certain, it is just possible that -sk and -cil are combinations of two original plural suffixes *-c and *-k and *-c and -l respectively. *-c might have been the plural suffix in such nouns originally and *-k or *-l have been tacked on to them because of the fact that the latter were more predominant in the respective subgroups". But, it must be mentioned at this point that "The plural in Dr(avidian) nouns and pronouns is not an obligatory category (except for most C(entral) Dr(avidian) languages which have developed it later.)" (Zvelebil:1990:21-22). So, one can ask, this being the case in Dravidian in general, why should the said languages have double plural markers at all? Thus, Subrahmanyam's explanation does not seem quite convincing and we will have to look for some other explanation.

In his analysis of Kolami data, Emeneau (1961:39, 42, 48-49) posits /-si-/ as a feminine gender marker and states, "All nouns with these suffixes denote female persons; the suffix -al, or rather the complex -a-l, with its allomorph -a-si, therefore means female person". Again, Tyler (1975:100), while talking about *Gondi*, observes that "....a more productive analysis would set up *-ac- as a separate feminine marker... before plural suffixes". Extending the same argument to Kui and Kuvi, it can be argued that /-ska/

consists of two morphemes like /-s-/ which is the feminine gender marker and /-ka/ which is the non-masculine plural marker. First, this /-s-/ is added to form the feminine stem, and then the stem takes the plural marker /-ka/. In other words, here /-s-/ should be treated as a part of the stem whereas /-ka/, as a suffix.

Regarding /-nga/, scholars like Andronov (1977), Caldwell (1974), Chetty (1984), Krishnamurti (1972), Winfield (1928) and Zvelebil (1977, 1990) hold the view that it is a variant of the plural marker /-ka/. Let us first discuss it with reference to *Kui*

Winfield (1928:16) states that "-ka is nasalized to -nga after roots ending in a vowel". Such a rule is highly unmotivated and there are many exceptions to it. Dash (1981:72) remarks that the distribution of /-ka/ and /-nga/ is not predictable. Chetty (1984:104-105) observes, "The plural suffix -nga generally occurs after disyllabic singular nouns which end with -a, -i, -e and -o. -ka generally occurs after disyllabic singular nouns which end with -u (which will be elided before it) and after trisyllabic singular nouns ending with -i (which will be elided before it)". Again, as there are plenty of counter examples to his observation, he concludes that "-ka and -nga are to be taken as morphologically conditioned (or lexically specific) allomorphs as they occur in the same phonological environment".

Then, a look at the *Kuvi* data given in Israel (1979:46-60) makes it clear that the most frequently used neuter plural markers are /-ka/ and /-yã/. "Broadly speaking, *-k appears after nouns ending in a consonant and *-nk after those ending in a vowel" in the South Central Dravidian languages (Subrahmanyam:1969b:98). The point to be mentioned here is that /-yã/ occurs only after vowels in *Kuvi* and this certainly strengthens our claim made above that it is a changed form of /-nga/. Now, I want to argue that /-nga/, like /-ska/, consists of two morphemes, viz. /-n-/ and /-ka/, and the latter becomes /-ga/ by assimilating the feature [+voice] of the preceding nasal. Further, this /-n-/ can be taken as the neuter marker in *Kui* and *Kuvi*.

In his discussion on *Kūbi*, Krishnamurti (1969a:225) recognises /-n(u)/ as a non-masculine plural marker. Similarly, Subrahmanyam (1968:177) states that the non-masculine plural morpheme in *Gonḍi* "... has two principal allomorphs, /-k/ and /-n/." Thus, the /-n-/ of /-nga/ in both *Kui* and *Kuvi* can be related to the /-n/ of *Kūbi* and *Gonḍi*. What is important here is that all non-masculine nouns, whether feminine or neuter, take the marker /-ka/ whereas the masculine nouns take /-rV/, i.e., /-ru/ in *Kui* and /-ri/ in *Kuvi*.

6. A pertinent question can be raised now: What was the cause behind the bifurcation of the non-masculine nouns into feminine and neuter in South

Central Dravidian? Subrahmanyam (1971:104) remarks that "The neuter plural suffix in Dravidian is an optional category in both S(outh) Dr(avidian) and N(orth) Dr(avidian). It is obligatory only in C(entral) Dr(avidian) and there it seems to be an innovation." If we can determine the factor that caused the use of the neuter plural suffix to be obligatory in South Central Dravidian, that will be the answer to the question about the bifurcation of the non-masculine into feminine and neuter.

On the separation of the feminine from the masculine and its merger with the non-human or neuter, Subrahmanyam (1969a:109) states "... the Central Dravidian languages other than Telugu separated the feminine category from the masculine and joined it with the non-human in the plural on the analogy of the distinction - masculine *versus* non-masculine that was already existing in the singular." But, our concern is to account for the emergence of the feminine plural marker in the group of languages under discussion. There seem to be two possibilities. First, it is the result of an effort to make a distinction between the feminine and the non-human which originally belonged to one class. It should be mentioned here that the Neo Indo-Aryan (henceforth NIA) languages spoken in India use separate plural markers for neuter nouns. For example, in Oriya, the most frequently used plural marker for non-human and inanimate nouns is /-gura/ whereas /-mane/ is used for the epicene or human nouns. The following instances will be illustrative:

<i>kālamāgura/*kālamāmane</i>	'pens'
<i>gāchāgura/*gāchāmane</i>	'trees'
<i>māsagura/*māsamane</i>	'mosquitoes'
<i>kukurāgura/kukurāmane</i>	'dogs'
??? <i>lokāgura/lokāmane</i>	'men'
??? <i>jhiāgura/jhiāmane</i>	'girls'

Sometimes /-gura/ can be used with human nouns in Oriya. But, its use expresses a strong pejorative sense for those human nouns to which it is added. In other words, it can be said that human nouns are demoted to the non-human category if /-gura/ is added to them. For this reason, /-gura/ should be taken as the non-human plural marker in Oriya. So, it is quite possible that the South Central Dravidian languages have made the neuter plural markers compulsory under the influence of the surrounding NIA languages like Oriya. Secondly, there was a pressure to retain a particular Dravidian structure and hence, the introduction of the said marker became necessary. This point needs elaboration. Krishnamurti (1972: 171-172), while discussing the phonological structures of the Proto-Dravidian verb-roots, observes that two main types of verb-roots can be established for Proto-Dravidian and each type has sub-types like the following:

Type	Sub-type	Possible number of occurrences
I	(a) (C) \bar{V}	44
	(b) (C) $\bar{V}C$	760
II	(a) (C) \bar{V}	44
	(b) (C) $\bar{V}C$	760

This statistics clearly indicates that the verb-roots in Proto-Dravidian were predominantly consonant-ending. The same is true about the nouns also. Thus, Zevebil (1990:13-14) rightly observes, "The Dravidian root has generally been recognised as being monosyllabic of the pattern (C) \bar{V} (C)" and "The typical stem be C \bar{V} C, too, but is commonly rather disyllabic and triconsonantal, CVC-VC." It is significant here to note that whether a root or a stem, it is normally consonant-ending in Dravidian. The second point to be mentioned here is that the South Central Dravidian group is quite conservative and that is why it retains many archaic characteristics. The following remark of Zvelebil (1977:11) lends support to this view: "The problem of reconstruction of the situation in P(roto) Dr(avidian) has not yet been satisfactorily solved. Final opinion has not been expressed either, whether P(roto) Dr(avidian) has the category of gender at all. It seems, however, that the situation as reflected by the *Kui-Kuvi* sub-family (plus *Gonḍi Konḍa Kōl(ami) Pā(rji)*) may be considered as the preservation of the original state of affairs." Emeneau also considers himself to be ".....on the right track as far as the archaic nature of *Kui-Kuvi* is concerned,...." (Zvelebil:1977:53, f.n. 13).

The important points here are that roots and stems in Proto-Dravidian are predominantly consonant-ending, and *Kui* and *Kuvi* are quite conservative languages. Given these facts, it is natural to hope that *Kui* and *Kuvi* would make an effort to retain the said Proto-Dravidian characteristic in its roots and stems. For this reason, I believe, whenever there is a non-masculine noun ending in a vowel, first a consonant, which is either /-s-/ or /-n-/ in this case, is added to it so that it conforms to the Proto-Dravidian stem-structure and after that, it takes the plural marker /-ka/.

7. To sum up, the main points discussed in this paper are the following: (1) *Kui* and *Kuvi* have only two plural markers, viz. /-rV/ for the masculine and /-ka/ for the non-masculine stems. (2) /-s-/ and /-n-/ should be treated as the feminine and the neuter markers respectively in these languages. (3) These markers developed most probably under the influence of the surrounding NIA languages which employ separate plural markers for the neuter nouns and also due to a pressure to retain the archaic consonant-ending Proto-Dravidian stem-structure in these languages which are quite conservative by nature.

COLOPHON

I am grateful to Professors P.S. Subrahmanyam and Kamil V. Zvelebil for their comments and suggestions on an earlier draft of this paper.

BIBLIOGRAPHY

- Andronov, M.S. 1977. *Dravidian Languages*. Vijayawada: Visalandhra Publishing House.
- Caldwell, R. 1974. *A Comparative Grammar of the Dravidian or South Indian Family of Languages*. Madras: University of Madras. (First published in 1856).
- Chetty, N.T. 1984. *Kuttia Dialect of Kui*. Unpublished M.Phil. dissertation. Hyderabad: Osmania University.
- Dash, G. 1981. *Structure of Kui: a Tribal Language of Orissa*. Unpublished Ph.D. dissertation. Santiniketan: Visva-Bharati.
- Emeneau, M.B. 1961. *Kolami: a Dravidian*. Annamalainagar: Annamalai University. (First published in 1955).
- Grierson, G.A. (ed. and comp.). 1967. *Linguistic Survey of India, Vol.IV*. Delhi: Motilal Banarsidass.
- Israel, M. 1979. *A Grammar of the Kuvi Language*. Trivandrum: Dravidian Linguistics Association.
- Krishnamurti, Bh. 1972. *Telugu Verbal Bases*. Delhi: Motilal Banarsidass. (First published in 1961).
-1969a. *Koṇḍa or Kūbi a Dravidian Language*. Hyderabad: Tribal Cultural Research and Training Institute.
-1969b. "Comparative Dravidian Studies". In: T.A. Sebeok (ed.), *Current Trends in Linguistics, Vol.5*, pp. 309-333. The Hague. Paris: Mouton.
-1975. "Gender and Number in Proto-Dravidian". *International Journal of Dravidian Linguistics* 4, 328-350.
- Shanmugam, S.V. 1972. "Gender-Number Sub-categorization in Dravidian". In S. Agesthalingam and S.V. Shanmugam (eds.), *Third Seminar on Dravidian Linguistics*, pp. 383-415. Annamalainagar: Annamalai University.
- Subrahmanyam, P.S. 1968. *A Descriptive Grammar of Gondi*. Annamalainagar: Annamalai University.
-1969a. "The Central Dravidian Languages". In S. Agesthalingam and N.K. Raja (eds.), *Dravidian Languages (Seminar Papers)*, pp. 107-134, Annamalainagar: Annamalai University.
-1969b. "The Gender and Number Categories in Dravidian". *Journal of the Annamalai University (Humanities)*. Vol.XXVI, 79-100.

.....1971. "The Characteristic Features of Dravidian Languages". In H.S. Biligiri (ed.), *Papers and Talks*, pp. 103-105. Mysore: Central Institute of Indian Languages.

Tyler, S.A. 1975. "Gondi h. With Some Notes on Number and Gender in Proto-Central Dravidian". In: H.F. Schiffman and C.A. Eastman (eds.), *Dravidian Phonological Systems*. pp. 86-113. Seattle: University of Washington.

Winsfield, W.W. 1928. *A Grammar of the Kui Language*. Calcutta: Asiatic Society of Bengal.

Zvelebil, K.V. 1977. *A Sketch of Comparative Dravidian Morphology (Part One)*. The Hague. Paris, New York: Mouton.

.....1990. *Dravidian Linguistics: an Introduction*. Pondicherry: Pondicherry Institute of Linguistics and Culture.

THE IMPERSONAL PASSIVE IN KANNADA

ANURADHA SUDHARSAN
Tribhuvan University, Kathmandu

A syntactic analysis of two impersonal constructions in contemporary Kannada suggests that they are not impersonal passives as such, but are, in fact, active sentences without an overt logical subject. It further suggests that the surface constraints of these sentences can be best explained if they are analysed as biclausal constructions, consisting of an infinitival complement and a copula with 3rd. person singular, neuter ending. In terms of *Relational Grammar*, they are monostratal constructions involving no relation, changing the rule whatsoever. In fact, the biclausal analysis argues against the spontaneous demotion analysis proposed in Sridhar (1979).

The subject-constraint of these sentences is explained in terms of the principles of GB theory and this is further supported by a historical account of these sentences. Interestingly enough, these sentences are shown to be of English in origin; one is supposed to be a translation of a biclausal *it* construction in English and the other a direct offshoot of another defective construction attested in 18th and 19th century popular Kannada prose. This defective construction is supposed to be an intermediate structure which developed as an analogical extension of the biclausal translated structure.

1. The discussion in this paper focuses two impersonal or passive-like constructions frequently used in contemporary Kannada. The sentence type illustrated in (1) is generally called an impersonal passive and has been discussed in Sridhar (1979). In its verbal morphology and syntactic constraints, sentence (2) is identical with (1), although historically they are not parallel constructions. According to the analysis I am going to present here, (1) originated from (2) via a process of analogical extension. This second sentence has not been discussed in Sridhar (1979) or elsewhere in the literature.

1. *Šāleyalli-makkalige-ondū manarañjane*
school.loc-children-dat-an entertainment
Kāryakramavannu-ērpadiśalāgide
programme-acc-arrange(caus). inf-become-is 3rd.sg.neut.
An entertainment programme has been arranged for the children in the school.

2. [*Vēda Kāladalli- Strīyarige-bahaḷavāgi*
 [Vedic times.loc-Women.dat-very much
Swātantryavittendu] *hēlalāgide*
 freedom-is-past-comp] Say inf-become-is 3rd sg.neut.

It has been said that Women enjoyed a lot of freedom during the Vedic times.

The sentence type illustrated in (1) has the following characteristics. It is exclusively a transitive verb construction but does not express a logical subject either as a nominative NP or even as an instrumental NP as shown by the ungrammatical (3a) below. It has a direct object NP with the accusative ending *-annu*. However, with a causative verb, a mediary agent can be expressed as in (3b).

- 3a. **dāktarinda/dāktaru-rōgige-aṣadiyannu-koḍalāgide*.
 doctor-inst/doctor.nom.-medicine.acc.-give.inf.is.....
- 3b. *Śāleyalli-tōṭada kelsasavannu- makkaḷindaṭē-*
 School.loc.-gardening.acc.-children.Inst.alone
mādisalāguttade
 get done-inf.-become-pres.3rd sg.neut.
 In the school, (they) have gardening done by the children alone.

The verbal morphology in this construction is generally supposed to be periphrastic. The transitive verb with the infinitival suffix *-al* is followed by the copula *-agu* 'to become, to be possible,' with a 3rd person, singular, neuter, inflection. As shown in (2), frequently an *endu*-clause can occur as the complement of the transitive verb, instead of an accusative NP: In other respects (2) is identical with (1).

In the following sections, I will attempt to present a detailed analysis of these constructions and explain their surface constraints and subsequently, give further support for my observations by way of a historical account of these constructions. Section 2 considers the spontaneous demotion analysis of (1) proposed in Sridhar (1979). Section 2.1 proposes a biclausal analysis of (1) and (2) -sentence (2) is, in fact, crucial to our analysis of (1), as we will see later. Section 2.2 looks at the history of (1) and (2) and argues that they are not native to Kannada, in that they developed in Kannada in a Kannada-English bilingual context. The discussion centres around the argument that these constructions originated in the Kannada translation of a biclausal *it* -construction in English. Section 3 concludes our discussion.

2.0. The Spontaneous Demotion Analysis: We will now consider Sridhar's analysis of sentence (1). Sridhar calls this sentence an impersonal passive. He accounts for the nonoccurrence of an overt logical subject by postulating a rule of spontaneous demotion by which the subject of a transitive verb is demoted to an oblique position or is completely deleted

without the corresponding advancement of the direct object to subject position (cf. Comrie:1977). According to Comrie (1977) deletion is also a form of demotion. Sridhar further argues that the subject although null, is present underlyingly; and he argues for its presence by showing that it is available for a number of transformations like Equi-Np deletion, Co-referential subject deletion etc., which refer to the notion of 'Subject' in Kannada. Sridhar's analysis infact argues in favour of Comrie's claim (Comrie:1977) that the Motivated Chomage Law proposed in Perlmutter and Postal (1977) rules out impersonal passives in many languages, and hence a rule of spontaneous demotion should be postulated as part of UG to account for impersonal constructions in many languages. Sridhar further suggests that the rule of spontaneous demotion (or deletion) applies after all other transformations have applied. By way of illustration, let's consider the application of Equi-Np deletion. The examples 4-6 are all from Sridhar (1979).

4. *avaru-sainikarannu-vāpassu kareyalu-oppidaru*
they- Soldiers acc.-back call inf.-agreed 3rd pl.
They agreed to call the soldiers back.
5. *sainikarannu-vāpassu kareyalu-oppalāyitu*
Soldiers-acc.-back call inf.-agree inf.become 3rd sg.neut.
It was agreed to call the soldiers back.
6. **avaru sainikarannu-vāpassu.kareyalu-oppalāyitu*
they-soldiers acc.-back call inf.-agree became

In (4), which is not an impersonal sentence, the subject of the matrix clause and the subject of the infinitival clause are coreferential; and the subject of the infinitival clause is, according to Sridhar, deleted by the Equi-NP rule. In the impersonal sentence (5), Equi-NP rule applies first, and then the spontaneous deletion applies, deleting the matrix subject. In (6) however, the spontaneous deletion rule has not applied subsequent to the Equi-NP deletion. Hence the ungrammaticality of the sentence. What this implies is that the underlying structure for (5) is (6), which has a specific nominative subject, which has to be deleted to derive the impersonal sentence.

Sridhar, however, goes on to consider an alternative analysis of (5) and (6). He suggests that we could assume null subjects for both the matrix and the complement clauses which would still meet the identity requirement for (6), and this analysis would not be empirically distinguishable from the spontaneous deletion analysis. The null-subject analysis, however, Sridhar argues, does not account for the ungrammaticality of (6), and also for the gap in the application of Equi-NP; besides, it ignores the fact that transitive verbs select their subject. Hence, Sridhar concludes, that only a spontaneous deletion analysis can account for both (5) and (6).

The problem with Sridhar's analysis is that it does not take into account the question about why the underlying subject is deleted instead of being demoted to an oblique position. In other words, he does not consider the question about what prevents this sentence from expressing a logical subject even as an instrumental NP? The rule of spontaneous demotion does allow the occurrence of an agent in the oblique position.

Besides, Sridhar's analysis does not seem to consider the problem of determining the subject of the copula *āgu*. It simply says the verbal morphology is invariant. In fact, one of the factors that motivates Comrie to postulate a spontaneous demotion rule for the Welsh impersonal passive is the absence of 3rd person, singular morphology in the construction. Some of these questions can possibly be answered within the paradigm of the GB framework. Let's now consider an alternative analysis which attempts to answer some of the above questions in terms of the principles of the Case theory and Theta theory.

2.1. The Biclausal Analysis: The problem with Sridhar's analysis lies mainly in the assumption that the construction in question is an impersonal passive and that it is derived by a transformational rule and also the theoretical framework within which he tries to explain the syntactic constraints of the construction. Rules like Equi-NP deletion which he uses to account for (5) are no longer part of the current GB theory. It is my contention that this construction is not an impersonal passive as such, but an active biclausal sentence, involving no transformation at all. It is impersonal in that it lacks an overt logical subject. Its surface constraints can be better explained by analysing it as a biclausal construction consisting of an infinitival clause functioning as the complement of the copula *āgu*. The infinitival clause is an active one in which the transitive verb has a direct object, with accusative marking (and has an indirect object too when it is double transitive). There is no passive morphology in the underlying structure to trigger a transformational rule to turn this active clause into a passive one- infact Kannada is devoid of a distinct passive morphology. The infinitival clause being untensed and hence its subject position being ungoverned cannot have an overt subject. In Kannada, agreement morphology is not necessary for the occurrence of a nominative subject. For instance, the so called participial relative clause, the absolute clause and the gerundive (nominalised) clause which do show tense variation, but which lack agreement inflection, can have an overt nominative subject as illustrated in (7), (8) and (9) respectively.

7. *Narēndranu hareyuttiruva/haredidda*

Narendra.nom tear away pres.prog/tear.past perf.

Pustakavu-namma laibrarī du

Book-our library's (is)

The book which Narendra is tearing away belongs to our library.

8. *maguvu-niddemāḍalu/niddemāḍuttiralu*
child-nom.-sleep.nonpast.inf./sleep prest.prog.
Sīta-aṅgaḍige hōḍalu
Sita shop to went 3rd sg.fem.
When the child was sleeping, Sita went to the shop.
9. *dinēṣānu-dinālu-gī tala manege hōguvudu/hōguttiruvudu*
Dinesha-nom.-everyday Gita's house-go.ger./go.pres.prog.ger.
namma tandege-istavilla
Our father dat.-liking-is not
My father does not like Dinesha going to Gita's house everyday.

On the other hand, the infinitival (-*al* and -*udakke*) clauses and adverbial participial clause which are not tensed, cannot have an overt nominative subject. So the infinitival clause in the impersonal construction is devoid of an overt nominative subject, but can have a PRO as its subject. This PRO is arbitrary since it is not coindexed with any other specific NP within the construction. The transitive verb assigns the agent θ role to PRO; hence there can be no other agent NP in this construction. Moreover, a transitive verb in Kannanda does not assign an agent θ role to an instrumental NP. A causative verb, however, can have a mediary agent NP with instrumental case marking. And a dative NP gets only the recipient theta role from a double transitive verb.

10. *PRO-arb.....NP acc....v + al (U)*
āgu + Tns + 3rd.sg.neut.

The infinitival suffix -*al* generally has an enunciative -*u* when it is not attached to any thing following it; but in the impersonal construction, the copula *āgu* occurs attached to the preceding suffix -*al* and so there is no -*u* after -*al* in this construction¹.

The infinitival clause in this construction is analysed as the complement of the copula -*āgu*-- in which case we assume an empty expletive as the subject of the copula, since Kannada is a null-subject language, or it can be analysed as the subject of the copula². These two

1. A rare instance of the impersonal construction in which the suffix -*al* occurs detached from the copula *āgu* is found in M.A. Ramanuja Iyengar's Kavi Kārya Praśamse (1912)

i. [Ondu Indriyada-sukhavannu-berondu Indriyada-
one sense organ's-pleasure-acc.-another sense organ's
mulaka-tillisalu]-aguvudilla
through-convey-inf-become (possible).pres.neg.
It is not possible to convey the pleasure experienced by one sense organ through another sense organ.

In this example, as we see, there is no overt agent. The transitive verb *till* 'to convey', has a direct object *sukhavannu* 'pleasure'. This sentence has all the features of the modern impersonal sentence shown in (1).

2. This observation is based on the intuitions of the scholars with whom I discussed this paper.

analyses are not empirically distinguishable since the subject is null and the infinitival clause always occurs preverbally and the subject before the infinitival clause.

As we see, the foregoing analysis does attempt to explain the gap in the distribution of nominative subjects and particularly of subjects of transitive verbs. In fact, I arrived at this analysis via a historical study and subsequently looked at it in the light of the GB theory.

2.2. Now, turning to the origin and development of this construction, we do not find any structure corresponding to this in Old Kannada or Middle Kannada. In fact, we do not come upon any similar construction even in the 17th and 18th century prose works like *Cāmsrājōktivilāsa*, *Punyāśrava*, *Hālāsyamahātme* etc., which are written in classical or traditional style. However, sentences of the type shown in (2) do appear in the 19th century 'popular' prose. And *Veerārājēndrawodeyar's Rājēndranāme*, a 19th century historical work which was later translated into English by Lieut. R. Abercrombie, is written in popular prose. It is simple and informal in its style and is full of Hindustani words. In the words of Herman Moegling, it is representative of "the Canarese language in its present phasis". And, Srinivasa Havanur (1974) is also of the opinion that 19th century heralds modern Kannada and that *Rājēndranāme* is the first (historical) work to appear in modern Kannada. Now, consider the following sentences (11) and (12) from *Rājēndranāme* and R. Narasimhacarya's *Karnāṭaka Kavicarite*, vol 2, respectively.

11. "....." *endu-bareyalāyitu*
that (thus) write-inf.-become-is 3rd-sg.neut.
Thus it was written.
12. [*ivarellā-uddāma kavigaḷendu*] *hēḷalāguvudilla*
all these (people)-great poets-that say.inf.-become-neg.
It is not possible to say that these are all great poets.

These sentences lack an overt logical subject and are identical with (2) in all respects and so would be perfectly acceptable in modern Kannada. We, however, come upon another kind of construction in the 19th century prose whose verbal morphology is identical with that of (1) and (2) but which differs significantly from the latter in other respects. Here are some instances of this construction from *Rājēndranāme*.

13. *hindu rājyadalli-halavu bage-balāḍyara saṅgada*
Hindu kingdom-loc.-several ways-powers with-
dōstiyāgi-tāvu-prayatna māḍiddu-
having made friends with- you.nom.-efforts make.past-ger.nom.
kelavu divasa āguvudakke muñce-namage-tiḷiyalāyitu
some days before-we.dat.-know.inf.become.3rd.sg.neut.
We came to know a few days before that you had made efforts to
engage yourself with strong powers in various negotiations.

In this sentence, the gerundive *udu*-clause, marked nominative, is the surface subject and the dative NP *namage* is the agent of *tīli* 'to know'. This construction would be considered ungrammatical in modern Kannada since it violates all the constraints that go with its verbal morphology. This construction type is often found in the early 20th century works also. Consider again sentences (14), (15), and (16) taken from M.A. Ramanuja Iyengar's *Kavi Kārya Praśamse*, Mudaveedu Krishna Rao's *Kannadavannu-ddharisida Punya Puruṣaru* and Masti Venkatesha Iyengar's *Thākūr Kaviyarara Gitāñjali*, respectively.

14. *avanannu-nōduttiddavara-nōduttiddavara-manōbhāvavanu-avaninda*
 he.acc.were seeing.rel.-feelings.acc.-he.inst.
vivarisalādite?
 describe.inf.-become.is.3rd.sg.neut.int.
 Is it possible for him to describe the feelings of those who were looking at him?

15. *alliya varegū-kannada śālegalalli-sikṣakarū-*
 until then-Kannada schools.loc.-teachers.nom.
parīkṣakarū-bahuśaḥ-marāt higarē-iruttiddaremba
 examiners nom.probably-Maharashtriyans-were that (rel.)
saṅgatiyu-āgina kālada-varīṣṭhādikārigaḷa-
 news-nom.-of that time-senior officers'
varadiyallē-nirdēśisalāgide.
 reports.loc.-point out.inf.-become.is.3rd.sg.neut.
 The news that until then the teachers and examiners in Kannada schools were all Maharashtriyans is pointed out in the reports of the senior officers of that time.

16. *kannadigaru-yārādarū-baṅgālī bhāseyannu kalitu-*
 Kannada people-anybody-Bengali language. acc.-having learnt
ā gi-tikegalannu-Kannadadalli-bareyuvavaregū
 those songs.acc.-in Kannada-until (they) write
Kannada nādinavaru-thākurara-i kāvyagala ruciyanu
 people of Kannada country.nom.-Tagore's-these poems' savour.acc.
tīliyalāguvudilla
 know.inf.become.pres-neg.
 Until some Kannada person learns the Bengali language and writes Tagore's songs in Kannada, Kannada people will not be able to appreciate the savor (quality) of these songs.

Sentences (14) and (16) have an overt agent, and sentence (15) has a surface nominative subject, which is actually the direct object of the verb *nirdēśisu* 'point out'. Each of the above sentences would be unacceptable in contemporary Kannada. One interesting aspect of this construction is that it is inconsistent as to how it expresses the agent. Sometimes the agent is expressed as an instrumental NP as in (14), sometimes as a dative NP as in (13) and sometimes as a nominative NP as in (16). Frequently, the direct object itself occurs as the surface nominative subject as in (15). Only when

there is an *endu*-clause, the agent or the surface subject is consistently absent as in (11-12).

How do we explain the behaviour of these two types of constructions possessing the same verbal morphology? One type, occurring with *endu*-clause, as shown in (11-12), keeps all the constraints while the other type shown in (13-16), which has the same verbal morphology violates these constraints. A possible explanation for this would be to hypothesise in the following way regarding the origin of these constructions. The sentence type containing an *endu*-clause is most probably a direct translation of a biclausal *it*-construction in English, illustrated below.

17. It is possible/difficult to say that..... This construction is quite frequently found in the 19th century expository or analytical English prose used by Indians and native speakers of English as well. This sentence is impersonal in that it lacks an overt logical subject. In this sentence, the infinitival and the *that*-clauses occur as verbal complements. In Kannada, the infinitival clause and the *endu*-clause are, in fact, instances of verbal complementation. Only nominalised and hence case-marked clauses can occur in subject or object NP positions. Since the infinitival and the *endu*-clauses are neither nominalised nor case-marked, they cannot occur in subject position. So these two clauses seem to share some similarities with the infinitival and the *that* clauses in English. Since Kannada is a null-subject language, it does not have an overt expletive subject corresponding to the *it* in English. This expletive element can, however, be recovered from the verbal morphology in the impersonal construction. In Kannada, the copula *-āgu* can mean either 'to become', or 'to be possible', while the copula *iru* 'to be' which corresponds to the English 'to be' does not have these meanings. Hence, the impersonal construction uses *-āgu* instead of *iru*.

Another feature of the above English sentence is that the infinitival verb is a transitive one with a *that*-clause as its complement. Thus we can establish neat correspondences between the English sentence and the Kannada impersonal construction (cf. eg.2).

As for sentences (13-15), they do not show such neat correspondences with the above English sentence or any other similar sentence in English. Nor are they easily traceable to any older construction in Kannada. However, since their verbal morphology is identical with that of (1) and (2), we can again hypothesise that these sentences are the result of an extension or an (over) generalisation of the use of the impersonal verb -as used in the biclausal translated structure- to NPs also and that the modern impersonal sentence illustrated in(1) is a direct offshoot of this defective construction and that the accusative-NP constraint developed

when this construction was reanalysed or restructured into an acceptable Kannada sentence. When the biclausal impersonal construction gained ground in popular usage, it was perhaps perceived as a monoclausal sentence, since the copula *āgu* always occurred, attached to the preceding infinitival suffix *-at*, both in pronunciation and orthography. This must have prompted Kannada speakers to generalise the use of the impersonal verb to NPs also. The resulting sentence was ungrammatical because of the 3rd person, singular neuter morphology of the copula. Gradually, this sentence must have fallen in line with the other biclausal construction and restructured itself accordingly. Once the sentence was reanalysed as biclausal, the overt logical subject had to disappear from the untensed infinitival clauses: or, the surface nominative subject in sentences like (15) had to be reinterpreted as an accusative NP. Further, it could be argued that syntactically the infinitival clause was reinterpreted as the complement of the copula *āgu*.

3.0. The foregoing analysis is rather simple and straightforward in that it accounts for the surface constraints of the impersonal sentence without postulating any transformational rule at all. In addition, the historical account ties in well with the syntactic analysis and also offers a reasonable explanation for the development of the subject-constraint and the transitive verb constraint in this construction. In fact, the transitive verb constraint is very important in explaining why this sentence cannot have even a dative subject. In sentences like the following, the copula *āgu* has a dative subject.

18. *nimage-āgadeiruvudu-nanage-āguttadeyē?*
 you-dat.-which is not possible-ger.-I-dat.-is it possible
 Is that which is not possible for you, possible for me?

This type of sentence does occur in Middle Kannada and probably in Old Kannada also. As we know, *āgu* can have an infinitival complement here. In fact, in *Vaddārādhane* (10th century), there are sentences in which the copula *āgu* occurs with a dative subject and an infinitival complement. Consider the following rhetorical question from *Vaddārādhane*.

19. *r̥ṣiyaram-nimage paḷiyalum/tiraskarisi nudiyalum akkumē?*
 r̥ṣis-acc.-you.dat..talk contemptuously of-is possible?
 Is it possible for you to talk contemptuously of the great r̥ṣis?

The infinitival verb in this example is transitive. But it is possible to think of a sentence of this type with an intransitive verb also, although we do not come upon such sentences in *Vaddārādhane*. But the impersonal construction cannot have a dative subject since the verb can be double transitive also, in which case the dative NP would get the indirect object interpretation and not the subject interpretation, as shown in (20)

20. *rōgige-ausadiyannu-kodalāgide*

Patient.dat.-medicine.acc.-give.inf.-become.is.3rd.sg.neut.

Medicine has been given to the patient.

So, a transitive verb assigns only a recipient θ role to the dative NP but not an agent θ role. Probably, the development of the impersonal construction had its impact on sentences like (18) in Kannada. In contemporary Kannada, sentences like (18) cannot have an infinitival complement ending in *-al*, but can have another type of infinitival which ends in *-udakke*, as illustrated in (21).

21. *nanage-mīrāṇnu-bhēṭimādōdakke-āgalilla*I-dat-*Mīra*.acc.-to meet-become-past.neg.It was not possible for me to meet *Mīra* (I could not meet *Mīra*).

All these observations seem to integrate well within the account I have proposed in this paper.

[This paper was presented at the 13th Annual Conference of the Linguistic Society of Nepal held in Kathmandu on November 26 and 27, 1992.]

BIBLIOGRAPHY

- Chomsky, N. 1986. *Knowledge of Language: Its Nature, Origin and Uses*. New York: Praeger.
- Comrie, Bernard. 1977. "In Defense of Spontaneous Demotion". In Peter Cole and Jerry Saddock (eds), *Syntax and Semantics*, 8: New York: Academic Press.
- Gai, G.S. 1946. *A Historical Grammar of Kannada*. Poona: Deccan College.
- Havanur, Srinivasa. 1974. *Hosagannadada Aruṇōdava*. Mysore: Institute of Kannada Studies, University of Mysore.
- Narasimhia, A.N. 1941. *A Grammar of The Oldest Kanarese Inscriptions*, Mysore: University of Mysore.
- Perlmutter, David and Paul Postal. 1977. "Toward a Universal Characterization of Passive". In Whistler et al.. (eds), *Proceedings of the Third Annual Meeting of the Berkeley Linguistic Society*. Berkeley: Linguistic Society.
- Shankara Bhatt, D.N. 1978. *Kannada Vākyagaḷu*. Poona: Deccan College.
- Sridhar, S.N. 1979. "New Evidence for Spontaneous Demotion". *International Journal of Dravidian Linguistics* 8:2, 312-322.

Note

**A CASE STUDY OF RELATIVIZATION AND
SYNTACTIC MARKEDNESS OF L₂ LEARNERS
OF MALAYALAM**

V. SARAT CHANDRAN NAIR
Central Institute of Indian Languages, Mysore

INTRODUCTION

Time and again, Chomsky (1981:6-8) has postulated that the ability to acquire a human language is genetically determined and that the learning task is endowed with Universal Grammar (UG) that comprises a finite set of innate principles that govern the emergence of a language¹. Now that he relates the theory of markedness with that of "core grammar" is particularly interesting to the study of problems in second language acquisition (SLA) and his statement, "to fix parameters of UG, providing a core grammar, guided perhaps by a structure of preferences and implicational relations among the parameters of core theory. If so, then considerations of markedness enter into the theory of core grammar", is equally valid.

The concept underlying the theory of markedness in linguistic terms is that more rules, additional features etc., would be required for the marked item as opposed to the unmarked item and in psycholinguistic terms, an item which requires more cognitive processing time is a marked member. According to Clark and Clark (1978:231),

If expression A can neutralize in meaning in contexts that the almost equivalent expression B cannot, then B is "more complex than A". Thus, more complex is equated to "more marked".

It could be presumed that the core rules are basic and would be learned/acquired easily with a minimum exposure to the language. On a comparison of the "theory of markedness" and "core grammar", we could say that the "unmarked" rules of a grammar would be the core and the marked rules would be the non-core. These assumptions have been

1 This paper is a revised version of the paper presented at the 16th All India Conference of Dravidian Linguists (January 13-16, 1988) at Trivandrum. I am thankful to the comments and criticisms of the participants of the seminar, as well as to Dr. E. Annamalai, Director, CIIL, who offered valuable comments at the stage of revising the paper. I am grateful to Sri. B.D. Jayaram, CIIL, for the help he has rendered in computing the results statistically.

validated in a number of developmental studies reported and are of particular importance (cf. Mazurkewisch:1985; Tarallo & Myhill:1983).

The present study shall focus on the implicational hierarchy proposed by (Keenan & Comrie:1977, 1979) and known as the Noun Phrase Accessibility Hierarchy (henceforth NPAH) and relate it to language learning proficiency. Maxwell (1979) has also discussed the strategies employed in relativization in different languages. The NPAH hypothesis has been used to predict successfully acquisition and accuracy orders for relative clauses in a second language (Gass:1979).

According to NPAH, there is a universal order in which NP positions (i.e., grammatical functions) can be relativized in natural languages. The order is as follows: Subject (SU), Direct object (DO), Indirect object (IO), Oblique object (OO), Genitive (G), Object of comparison (OC). Thus, SU is more accessible to relativization than DO, DO is more accessible than IO, and so on down the hierarchy upto G. This study has not included OC in Malayalam². According to this hypothesis, if a language can relativize IO, then it will relativize all other positions above it i.e., DO and SU, but not necessarily the other positions. This is interpreted as the implicational scale with SU at one end and OC at the other end. Thus, a sentence in Malayalam like 1, if embedded, can generate two structures as 1a and 1b.

1. *rāman marattilninnā vīṇu*
Raman tree-from fall-past
Raman fell from the tree
- 1a. *marattilninnā vīṇa rāman...*
tree-from fall-past-R.P. Raman
Raman who fell from the tree...
- 1b. *rāman vīṇa maram...*
Raman fall-past-R.P. tree
The tree from which Raman fell...

Obviously, (1b) is more difficult to be generated than (1a). In 1b, the head noun is from OO position, whereas in (1a), the head noun is from SU position. In (1b), the deletion of the case marker of the head noun is an additional rule, which would be required. So, more rules are required in the generation of (1b) than (1a). Thus, 1b is the marked form (M), and

2 The L2 learners of Malayalam studied in Nair (1984, 1986), had their mother tongues as Hindi, Oriya and Tamil respectively. L2 learners of Malayalam for the present study were 13 graduate teacher-trainees of Malayalam at the Southern Regional Language Centre, CIIL, Mysore, who had undergone a six-month intensive course in Malayalam and whose mother tongues were Oriya, Hindi, Kannada and Telugu respectively. The Malayalam-mother-tongue groups, treated as a control group, were deputed to the Centre to learn Tamil, Telugu and Kannada respectively. All the adult teachers were in the age group of 25-40 years.

(1a) is the unmarked form (UM), in the generation of relative clauses in Malayalam.

The scope of the problem becomes much more relevant because it was observed that a certain aspect of the target language (TL) is acquired easily than a related aspect of the TL, be it at the phonological, morphological, syntactic or semantic level of the TL in question. Thus, it could be presumed that the advanced alveolar trill is acquired more easily than the retracted alveolar trill or the accusative case is acquired earlier than the sociative case in second language acquisition (SLA) situation.

It was found in an earlier study (Nair:1986) irrespective of the learners' mother tongue, errors of a certain kind were common at all levels of the grammar. For example, the confusion at the sandhi level of words ending with /-a/ or /-am/, which causes the learners to produce significant IL forms like /* *marayil* / instead of /*marattil*/ 'in the tree' and /* *sadyattil* / instead of /*sadyayil*/ 'in the feast' etc. Similar common errors were found at the syntactic level too. To cite certain examples, the accusative case marker is assigned in the place of sociative case for communicate verbs as exemplified below,

2. /* *bassil kayarān addēham enne paraññu* /
 Bus-loc.M̄ get-in-Inf. M he I-Acc.M tell-past
 He told me to board the bus

instead of

- 2a. /*bassil kayarān addēham ennōt a paraññu* /
 bus-loc.M̄ board-Inf.M. he I-Soc.M̄ tell-past
 He told me to board the bus

Another pertinent error that L2 learner of Malayalam commits is that while trying to distinguish between genitive Noun Phrase and the NP, which is modified by the determiner with *ile constructions*. The learner produces either the genitive NP or the NP with locative case marker. On the other hand certain learners distinctly contrast -ile constructions³ and the Genitive NP, whereas for others it is in complementary distribution with the locative case. It seems, that in the case of learners, who produce the locative case, it is due to the phenomenon of neutralization. See below the errors identified.

3. * *ā Pradēsattinre malapradēsam*
 that place-Gen.M. hilly-place
 The hilly place of that place

3 Malayalam has a peculiar construction as opposed to other Dravidian languages, which is termed as -ile constructions. For example, /*kāttile āna*/ meaning 'Elephant of the forest' See Variar, K.M. Prabhakara (1979:98) for better elucidation.

3a. *ā pradēsattile malapradēsam*
that place-Loc. phr. hilly place

4. * *kātt-il āna*
Forest-Loc. C.M. elephant
The elephant of the forest

4a. *kāttile āna*
Forest-Loc.phr. elephant
The elephant of the forest

Deleting the locative case marker and using the noun to which the case marker should have been appended, not making a distinction between the use of intransitive Vs. transitive or transitive Vs. causative, using certain quantifiers wrongly, are some of the other common errors at the syntactic level, irrespective of the language backgrounds of the learner (for details, see Nair:1986:289-304).

Procedure

The NPAH hypothesis with regard to the acquisition of relative clauses in each position of L₂ learners of Malayalam, who had different mother tongues and the data were compared with 8 native speakers of Malayalam under the same set of conditions who produced relative clauses from the same set of basic sentences.

Elicitation material is presented in Appendix I. It consists of 12 sentences in which there are 12 occurrences in SU position, 4 in DO, 1 in IO, 6 in OO and 2 in G position. The basic sentences were read to two L₂ learners and two MT speakers and were asked to create RC structures for each sentence followed by the appropriate matrix sentence. The whole procedure was audio-taped and analysed. On the same day, the basic sentences were provided to both the groups and were asked to create as many RC structures as possible. It was found that there was high correlation between the recorded material and written material and hence the written material was taken as more authentic⁴ since the number of people tested were more. The interview for each person lasted 15 minutes when it was audio-taped and 30 minutes were allotted for writing for both the groups.

It was found that MT speakers also found the production of RC structures difficult in certain positions. MT speakers could not achieve a 100% success in the production of RC. MT group was taken as the control

4 It was observed that the actual IL of the learner could be ascertained more in an informal context than in a formal context.

group and L2 learners as the experimental group. The data were compared with the control group. Only correct RC structures were taken into account for tabulating the results. L2 learners often produce structures without deleting the case markers. For example, forms like (5) were produced for the given basic structures (5a) and (5b). (5c) should have been the correct structure.

5. * *rema ammayutekūṭe pōya ambalattil nān kaṇṭu*
I saw the temple to which Rema went with mother

5a. /*rema ammayutekūṭe ambalattil pōyi*/
Rema mother Soc.C.M. temple-Loc.C.M. go-past
Rema went with mother to the temple

5b. /*ā ambalam nān kaṇṭu*/
that temple I see-past
I saw that temple

5c. *rema ammayutekūṭe pōya ambalam nān kaṇṭu*
Rema mother-with go-R.P temple I see-past
I saw the temple to which Rema went with mother

The significant point to be observed is that why the learners retain the case marker in the head noun as /*ambalattil*/ in (5). It should have been /*ambalam*/ as in (5c). Is it an intermediary stage in the learning process?

The result of the performance of both groups were statistically compared as follows using 't' test.

1. Comparison between MT speaker and L2 learner at SU, DO and OO positions were made. This was to determine the proficiency or achievement of the learner in Malayalam with regard to relativization. SU, DO and OO positions were taken because IO and G had very low occurrences. It yielded the following result as shown in Table 3.

Table 1: MT speakers performance -RC formation

When the head NP is	SU	DO	IO	OO	G
Standard Score	1.0	0.33	0.08	0.5	0.16
MT ₁	0.83	0.25	0	0.33	0
MT ₂	0.83	0.33	0	0	0
MT ₃	0.83	0.33	0	0.25	0.08
MT ₄	0.75	0.33	0	0	0
MT ₅	1.0	0.25	0	0.25	0.08
MT ₆	0.58	0.25	0	0.41	0.08
MT ₇	0.75	0.33	0.08	0.5	0.08
MT ₈	0.83	0.33	0	0.33	0.08
Average	0.80	0.30	0.08	0.25	0.05

Table 2: L₂ learners performance-RC formation

When the head NP is	SU	DO	IO	OO	G
A [Or]	1.0	0.33	0	0.33	0
B [Or]	1.0	0.33	0.08	0.5	0
C [Or]	0.91	0.33	0.08	0.41	0.08
D [Or]	0.83	0.16	0	0	0.08
E [Or]	0.83	0.25	0	0	0.08
F [Or]	0.66	0.33	0	0	0.08
G [K]	0.83	0.33	0.08	0.41	0.08
H [K]	0.75	0.33	0	0.41	0.08
I [Te]	0.91	0.16	0	0.41	0.08
J [H]	0.91	0.16	0	0.08	0.08
K [H]	0.66	0.08	0	0	0
L [H]	0.83	0.16	0	0.16	0
M [H]	0.75	0	0	0	0

Or Oriya speaker learning Malayalam

K Kannada speaker learning Malayalam

Te Telugu speaker learning Malayalam

H Hindi speaker learning Malayalam

Table 3: 't' Values at each position between MT Speakers and L₂ learners

Positions	SU	DO	OO
't' values	0.710	1.792	0.940

To assess the implicational scale and NPAH hierarchy, the data were compared between SU, DO and OO positions for both MT speakers and L₂ learners. The percentage was computed as follows in Table 4 and 5. The 't' test applied is shown in Table 6. Figure I represents deviations in the implicational hierarchy at each position.

Table 4: MT speakers performance of RC formations-percentage

	SU	DO	IO	OO	G
MT ₁	83	75	0	66	0
MT ₂	83	100	0	0	0
MT ₃	83	100	0	50	50
MT ₄	75	100	0	0	0
MT ₅	100	75	0	50	50
MT ₆	58	75	0	83	50
MT ₇	75	100	100	100	50
MT ₈	83	100	0	66	50
Average	80	90.62	12.5	51.87	31.25

Table 5: L2 learners performance of RC formations-percentage

	SU	DO	IO	OO	G
A	100	100	0	66	0
B	100	100	100	100	0
C	91	100	100	83	0
D	83	50	0	0	50
E	83	75	0	0	50
F	66	100	0	0	0
G	83	100	100	83	50
H	75	100	0	83	50
I	91	50	0	16	50
J	91	25	0	0	0
K	66	25	0	0	0
L	83	50	0	33	0
M	75	0	0	0	0
Average	83.61	67.30	23.07	35.69	19.23

Figure 1: Percentage of relative clause generated by mother tongue speakers (MT) and L2 learner of Malayalam (L2) in each position

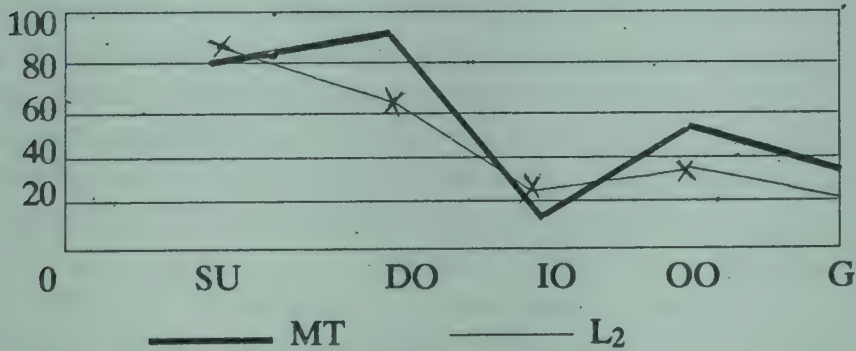


Table 6: ‘t’ values for the implicational scale between SU, DO and OO positions, for MT speaker and L2 learner

	SU > DO	DO > OO
MT	1.698	2.603
L2	1.774	4.20

The result shows that at SU, DO and OO positions, ‘t’ values in Table 3 are not significant implying that L2 learner has achieved native-like proficiency in regard to the production of RC structures. It is presumed that at IO and G positions also the result would have been too insignificant. The result shown in Table 6 indicates that for MT speakers, the production of RC construction at SU and DO positions is not significant whereas between DO and OO, the difference is significant,

which indicates that OO is more marked than DO and SU positions. On the other hand, for the L₂ learner between SU and DO positions, it is significant marginally and clearly significant between DO and OO positions.

The results obtained in this study support the general hypothesis that SLA progresses from unmarked positions in terms of the order in which the target language RC structures are mastered. The fact that the least proficient student is able to relativize only at the SU position, but the most proficient student is able to relativize at all other positions, proves this. The NPAH suggested by Keenan and Comrie (1977) is statistically validated as far as the acquisitional sequences of RC formation in Malayalam are found for L₂ learners with different mother tongue backgrounds. It also showed that L₂ learners have attained native-like proficiency in the production of RC structures, at three positions examined, viz, SU, DO, IO, at a later stage of learning. The hypothesis could still be further validated using bigger samples and at all positions of the NPAH.

However, it was found that MT speakers scored a higher percentage at the DO position than at SU position as shown in Table 4. This is intriguing. Is it due to 'focus', is a point to be pondered over? A possible explanation would be that MT speakers preferred to relativize NP at DO than at SU position. The percentage obtained in Table 5 indicates that L₂ learners clearly validates the markedness phenomenon in the order of acquisition of RC structures. Even for a poor learner like 'M' with Hindi as mother tongue, the hypothesis is validated. He could relativize only at SU positions and at no other positions. For him, SU being the least marked, which requires lesser rules, he has acquired that. The linguistic factors that would be bothering him with regard to the formation of RC structures would be

- (i) Movement of NPrel
- (ii) Deletion of case markers of the head noun
- (iii) Assigning particular marker for the verb of the embedded clause.

and hence, he generates RC structures like, **/oru maram muriccu Kṛṣṇan kōṭālikontə/*, while attempting to generate a structure like */Kṛṣṇan maram muricca kōṭāli/* 'The axe with which Krishnan cut the tree'. One reason for retention of case markers would be that, in Hindi, the mother tongue of the speaker, the case marker is retained in non-reduction RC type. It is surprising that the learner, 'I', with Telugu as mother tongue also generates RC structures without deletion of the case markers in the OO positions. Thus, a phenomenon which is not present in MT, is noticed in the IL. Retention of case markers may be one of the IL forms, and which is an

unmarked feature, and case deletion is a marked phenomenon. It is typologically valid since there are many languages, which retain case markers after relativization. However this paper does not go into the details of the strategies employed in relativization, by L2 learners and there is great scope in studying this area.

Originally, the markedness notion according to Roman Jakobson was conceived to be a binary notion. But, later developments showed that markedness could be extended to features and it is not merely a binary notion because in languages, there exist such phenomena as the notion of singular, dual, plural or masculine, feminine, neuter etc. Thus, markedness was revised and treated as distinctive features. See Andrews (1990) for details. Hence, the view taken in this article is that subject position is the least marked item and it is proved that the least proficient student is able to relativize only at the SU position that the most proficient student is able to relativize at all other positions. So 'SU' is the least marked in relation to all other positions.

The findings of this study give further evidence for the markedness hypothesis in SLA at the syntactic level. The implicational order in which NPAH is proposed is well founded with regard to the acquisition of relative clause by L2 learners of Malayalam with different mother tongue backgrounds. The hypothesis could be further validated using different samples and at all positions and in different language learning situations. similar study could also be done with regard to the acquisition of relative clauses in a non-formal context and compared with a formal situation so that we could generalise the hypothesis.

Bibliography

- Andrews, Edna. 1990. *Markdeness Theory, the Union of Asymmetry and semiosis in Languages*. Durham and London: Duke University Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.
- Clark, H and E. Clark. 1990. "Universals, relativity and language processing". In Greenberg J (ed). *Method and Theory. Universals of Human language*, Vol.1. Stanford University Press.
- Comrie, Bernard and E.L. Keenam. 1979. "Noun phrase Accessibility revisted." *Language*, Vol.55, No.3, Sept.1979.
- Eckman, F.R. 1977. "Markedness and Contrastive analysis Hypothesis" in *Language Learning*, Vol.27: p. 315-330.
- Gass, S. 1979. "Language Transfer and Universal Grammatical relations". *Language Learning*, Vol.29: p. 327-344.
- Keenan, Edward L and Bernard Comrie. 1977. "Noun phrase Accessibility Hierarchy and Universal Grammar". *Linguistic Inquiry*, Vol.8: p.63-99.

- Mazur Kewisch, Irene. 1985. "Syntactic Markedness and Language Acquisition". *Studies in second Language Acquisition*, Vol.7: p.15-37.
- Maxwell, Daniel.N. 1979. "Strategies of relativization and NP accessibility". *Language*, Vol.55, No.2.
- Nair, V. Saratchandran. 1986. "A Study of the errors of the Second Language learners of Malayalam. *Proceedings of the 13th All India Dravidian Linguists Conference*. 287-306. Thanjavur: Tamil University.
- Rutherford, William.E. 1982. "Markedness in Second Language Acquisition". *Language Learning*, Vol.32: p.85-108.
- Trallo. F & J. Myhill. 1983. "Interference and natural language processing in Second Language Acquisition". *Language Learning*, Vol.33: p.55-76.
- Variar, K.M. Prabhakara. 1979. *Studies in Malayalam Grammar*, Madras. University of Madras.

Appendix

Sentences which were tested for the hypothesis: Read the following basic sentences 1 to 12 and create as many relative clauses as possible. You are expected to create and write down the relative clause which comes to your mind first. Your first decision is which we want. A model was also demonstrated a, b & c of 1 to 12 are the various relative clauses to be generated.

BS 1. *rāman marattil ninnə vīṇu*

Ram fell from the tree

UM 1.a. *marattil ninnə vīṇa rāman* (SU)

'Raman who fell from the tree'

M 1.b. *rāman vīṇa maram.....*

The tree from which Raman fell (OO)

BS 2. *innale ī kasērayuṭe kāl oṭiṇṇu.*

'Yesterday the leg of this chair broke'

UM 2.a. *innale oṭiṇṇa ī kasērayuṭe kāl....*(SU)

'The leg of this chair which broke yesterday'

M 2.b. *innale kāl oṭiṇṇa ī kasēra*(G)

'The chair which broke its leg yesterday'

BS 3. *rema ammayuṭe kūṭe ambalattil pōyi*

'Rema went to the temple with mother'

UM 3.a. *ammayuṭe kūṭe ambalattil pōya rema...*(SU)

'Rema who went to the temple with mother'

M 3.b. *ammayuṭe kūṭe rema pōya ambalam* (OO)

'The temple to which Rema went with mother'

M 3.c. *remayuṭe kūṭe ambalattil pōya amma* (G).

'The mother who went with Rema to the temple'

BS 4. *kr̥ṣṇan̄ rāman̄ a oru pustakam koṭuttu*

‘Krishnan gave a book to Raman’

UM 4.a. *rāman̄ a oru pustakam koṭutta Kr̥ṣṇan̄.....(SU)*

‘Krishnan who gave a book to Raman’

M 4.b. *kr̥ṣṇan̄ rāman̄ a koṭutta pustakam...(DO).*

‘The book which Krishnan gave to Raman’.

M 4.c. *kr̥ṣṇan̄ āl oru pustakam koṭukkappetta raman̄(IO)*

‘Raman who was given a book by Krishnan’

BS 5. *rāghavan̄ rājuvin̄ōṭ a oru katha paraññu*

‘Raghavan who told a story to Raju’

UM 5.a. *rājuvin̄ōṭ a oru katha parañña rāghavan̄..(SU)*

‘Raghavan who told a story to Raju’

M 5.b. *rāghavan̄ rājuvin̄ōṭ a parañña katha...(DO)*

‘The story which Raghavan told Raju’

BS 6. *sīta oru pāṭṭi a pāṭi*

‘Sita sang a song’

UM 6.a. *pāṭṭu pāṭiya sīta..(SU)*

‘Sita who sang a song’

M 6.b. *sīta pāṭiya pāṭṭi a.....(DO)*

‘The song which Sita sang’

BS 7. *māṭēvan̄ vellattilēykkā āśayōṭe nōkkininnu*

‘Mateven was looking at the water with great desire’

UM 7.a. *vellattilēykkā āśayōṭe nōkki ninna māṭēvan̄ (SU)*

‘Mateven who was looking at the water with great desire’

M 7.b. *māṭēvan̄ āśayōṭe nōkki ninna vellam...(OO)*

‘The water which Mateven was looking with great desire’

BS 8. *kr̥ṣṇan̄ kōṭālikon̄ṭ a maram muriccu*

‘Krishnan cut the tree with an axe’

UM 8.a. *kōṭālikon̄ṭ a maram muriccakr̥ṣṇan̄....(S)...*

‘Krishnan who cut the tree with an axe’

M 8.b. *kr̥ṣṇan̄ kōṭālikon̄ṭ a muricca maram..(DO)*

‘The tree which Krishnan cut with an axe’

M 8.c. *kr̥ṣṇan̄ maram muricca kōṭāli (OO)...*

‘The axe with which Krishnan cut the tree’

BS 9. *pulakalēllām malakalilinn̄ a utbhavikkunnu*

‘All the rivers are originating from the mountains’

UM 9.a. *malakalilinn̄ a utbhavikkunna pulakal̄....*

‘The rivers which originate from the mountains’

M 9.b. *pulakal̄ utbhavikkunna malakal̄ (OC)....*

‘The mountains from which the rivers originate

BS 10. *maññakkil̄i kombilēykkā parannu*

‘The yellow bird flew towards the branch’

UM 10.a. *kombilēykkā paranna maññakkīli (SU)..*

‘The yellow bird which flew towards the branch’

M 10.b. *maññakkīli paranna kombā (OO)....*

‘The branch towards which the yellow bird flew’

BS 11. *kāttā vīśi*

‘The wind blew’

UM 11.a. *vīśiya kāttā (SU)...*

‘The wind which blew’

BS 12. *paṭṭi kuraccu*

‘The dog barked’

UM 12.a. *kuracca paṭṭi*

‘The dog which barked’

Abbreviations

BS	Basic Sentence	DO	Direct Object
UM	Unmarked	IO	Indirect Object
M	Marked	OO	Oblique Object
SU	Subject	G	Genitive
*	Errors		

Note

**AUXILIARY PREFERENCE AND PERFORMANCE
BEHAVIOUR OF PUNJABI SUBJECTS UNDER
CONTEXTUALIZED AND NON-CONTEXTUALIZED
SITUATIONS**

SETJIT KAUR
Punjabi University, Patiala

1. English Modal Auxiliaries: Meaning Form and Usages

Apart from being used as a medium of query and exchange of information, language also helps in communicating requests, offers, suggestions, wishes and intentions. A set of verbs that help to conduct the latter functions of language are called modals. English modals are **can, could, may, might, shall, should, will, would, must, ought to**¹.

English modals are semantically complex. 'What makes it so difficult to account for these words (.....modal auxiliaries...) is that their meaning has both a logical and a practical (or pragmatic) element. We can talk about them in terms of such logical notions as 'permission' and 'necessity,' but this done, we still have to consider ways in which these notions become remoulded by the psychological pressures which influence everyday communication between human beings: factors such as condescension, politeness, tact and irony" (Leech, 1971: 66-67).

Modal auxiliaries or **Secondary auxiliaries** (Palmer, 1965: 16) represent one of the three main functional forms of a class of words called verbs; the two other verb forms being full (or lexical) verbs and primary verbs (**BE, HAVE** and **DO**). Modal auxiliaries help to capture the **modality** i.e. "the manner in which the meaning of a clause is qualified so as to reflect the speaker's judgement of the likelihood of the proposition it expresses being true" (Quirk et al., 1985: 219). The nature of modality may be **intrinsic** (when some kind of intrinsic human control over such events as permission, obligation or volition exists) or **extrinsic** (when some kind of human judgement of events involving possibility, necessity and prediction exists). Modals have both of these uses, although there are areas of 'overlap' and neutrality between the intrinsic and extrinsic senses of a modal. For instance, the **will** in a sentence *I will see you tomorrow* can be said to combine the meanings of volition (determination, desire, willingness) and prediction.

Modal auxiliaries are related to mood and signal such hypothetical situations as may relate to possibility, necessity, ability, obligation,

intention, permission and assertion (Hakutani, 1972: 11). Modals are used to state an intention (*I will help you when I am ready*), express a wish (*I would like to inquire about it*), indicating importance (*you must do it at once*), or introducing what you are going to say (*I must admit it*). From the semantic point of view, thus, modal auxiliaries are often specialized towards the expressions of such speech acts as giving advice (**ought to**, **should**), making promises or threats (**will**), giving orders (**must**, **can**), etc. These auxiliaries differ from other verbs of English in one striking way, i.e. their occurrence in the four syntactic structures called **negation**², **inversion**³, **code**⁴, **emphatic affirmation** (Palmer, 1965:18).

The basic structural aspect of the modal auxiliaries is relatively straight forward. Modals are normally followed by the infinitive, which is bare (i.e., the base form of the verb alone) except with **used** and (usually) **ought** (where infinitive is followed by **to**) which are somewhat 'marginal' to the class of modals (Quirk et al. 1985: 127). Modal auxiliaries occur only in finite⁵ functions. They do not inflect in the case of present time reference. This means that there is no 's' form in the third person singular (**musts*, **cans*), and there is no '-ing' (present participle) or '-ed' form (past participle). They cannot occur in non-finite functions, i.e. as infinitives (**to may*) or present /past participles (**maying*, **mayed*) and as a consequence of this, they can occur only as first verb in the verb phrase.

Modals are usually neutral to time reference, i.e., the semantic categories of past, present, and future⁶. Past forms of modals can refer to past time also. However, when a modal is followed by an auxiliary verb **have** (modal + perfective) or **be** (modal + progressive) and a participle (present participle or past participle), the time reference gets marked. There are two exceptions, **will** and **shall**. Whereas **shall** always indicates a future event or situation (*I shall do what you suggested*), **will** usually indicates a future event or situation and, at times, refers to a present situation (*Will you give me a glass of water?*). Not only the present forms, but the past forms of the modal auxiliaries can be used to refer to present and future time (often with hypothetical or tentative meaning). Modal auxiliaries which do not have a distinct past form (e.g. **must**, **ought to**) can be used to refer to the past in indirect speech.

In passive structure, a modal is followed to **be** or **have been** and a past participle. A shift of voice in verb phrases may be accompanied by a shift in modal meaning. For instance, the modal **can** in the active sentence *He can do it* expresses ability, but in the passive sentence *It can not be done by him* expresses possibility.

Modals are often used to produce a particular effect, say, to get someone to behave in a particular way, to take a particular action, to accept an offer, or to give permission for something to be done. The choice

of a modal helps in identifying (a) the formality or informality of the situation, (b) the intensity of the relationship between the speaker and the listener, (c) the degree of politeness and persuasiveness and (d) the degree of importance attached to what is being said. Modals are used to explain the possibility of something happening or being done, to express degree of certainty about past, present and future situations and events, to say that something is permissible and to say that something is forbidden or unacceptable.

2. Punjabi Modals

Let us recall that English modals are one of the three main constituents of the verbal phrase: the two other verb forms being full (or lexical) verbs and primary verbs (BE, HAVE and DO). In Punjabi, modals form a sub-category of operator⁷, which, in turn, represents one of the three main constituents of the verbal phrase viz. main verb, operator and auxiliary verb. Punjabi modal sub-category has five members: /sak/, /cuk/, /das/, /ho/ and /kar/. If a member of other sub-categories of operator (primary, passive or progressive operator) co-occurs with modal operator in the structure of the verbal phrase it must precede it. These modals occur in the /dā/, /-ia/, /-u/, /-o/, /-e/ and /-egā/, forms with the exception of /ho/ which occurs in its /-dā/ and /-ia/ forms with exception of /ho/ which occurs in its /-dā/ and /-ia/ forms (Puar, 1990: 36). For instance, the Punjabi sentence

/ə ~ grezī bolṇī kādī vī sikkhī jā sākḏī h ∈/

(Speaking English can be learned any time.)

is an illustrative example of the use of modal operator /sak/ in its -da form (/sākḏī/); the sentence

/e' mākān tā ~ khārī diā jā cukkiā h ∈/

(This house has been bought.)

is an example of modal operator /cuk/ in its /-ia/ form (/cukkiā/). Modal operator /das/ in its /-da/ form (/dāssī de/) as in

/is hāthiār to ~ hor vī kə ~ m lāe jānde dāssī de hān/

(It is said that this weapon was used for various other purposes.),

modal operator /ho/ in its /-da/ form (/hu ~ da/) as in

/jədo ~ vī m ∈ ~ usde kà jā ~ dā ta ~ ó guṭkā l ∈ ke pāṭh kār riā hu ~ da sī/

(Whenever I visited his house he was found reciting prayers holding the holy book in his hand.)

and modal operator /kar/ in its /-da/ form (/kārda/) as in

/jāvānī vic me ~ kāsraṭ kār liā kārda: sã /

(When I was young I used to take exercise.)

are some other examples of the use of modal operator sub-category.

Modal operator sub-category (though the last sub-category among the operators) is always followed by the auxiliary sub-category of the verb, if represented. Unlike English modal auxiliaries which admit inversion in interrogative sentences, the members of Punjabi modal sub-category exhibit no change in interrogative sentences. All that is needed to form a question is an insertion of element /k̄/ in the initial position of the sentence.

3. Modal Auxiliaries: Performance Rate and Preference Patterns

Two sets of questions were administered to students to ascertain the level of correct use of modal auxiliaries and preference patterns, if any. The first set consists of seven non-contextualized interrogative sentences (S.1 to S.7, Table 1); five of which require the filling-in of modal auxiliaries in the initial position and two are direct quotes where fill-in entry forms a part of direct speech. The second set consists of nine sentences (C.1 to C.9, Table 2) that forms a part of a contextualized (Cloze) passage (Kaur, 1993: Chapter 7); five of which are interrogative sentences and four declarative sentences.

The average correct performance rate is manifold higher in the case of non-contextualized set (73 per cent) than in the contextualized set (28 per cent with exact entries and 38 per cent when appropriate entries are also considered). Thus, the performance rate is highly sensitive to the degree of choice available; the choice of modal auxiliaries is highly restricted in the contextualized setting.

In both the sets, the lowest preference is accorded to the use of **should**: its average correct performance rate in the non-contextualized set (in three sentences, S.1, S.2 & S.5) and in the contextualized set (two sentences, C.2 & C.7) is 11.92 per cent and 9.00 per cent respectively (Table 3). So is the case with **would**, the next least frequently used modal auxiliary; its average correct performance rate in the non-contextualized set (in four sentences, S.3, S.4, S.6 & S.7) and in the contextualized set (two sentences, C.6 & C.9) is 28.38 per cent and 14.12 per cent respectively (Table 3).

Its implication is that Punjabi students fail to display a good grasp in the use of **would** and **should**. These forms of modal auxiliaries are used to

express willingness in a tentative and polite way in situations related to offers, requests and suggestions.

However, the two sets differ in the priority assigned to the correct use of modals other than **would** and **should**. Let us first take the modal auxiliary **can**. Its use is admissible in all the seven non-contextualised sentences, but permissible only in three of the nine sentences of the contextualized set. Its average performance rate is substantially low (22.25 per cent) in the non-contextualized set where other alternatives are also admissible compared to the contextualized set (43.25 per cent) which does not permit the use of its substitutes (Table 3).

Modal Auxiliaries

Table 1: Frequency of the use of different modal auxiliaries in seven non-contextualized sentences by 400 students of 10 + 2 level from colleges affiliated to Punjabi University, Patiala, 1993

Sent no.	←Response→		←Correct modal auxiliaries→						Total
	Nil	Incorrect	will	would	can	may	should	shall	
S.1	5	38	0	0	125	166	12	54	400
S.2	5	75	0	0	74	108	46	92	400
S.3	5	86	141	64	104	0	0	0	400
S.4	3	74	72	141	110	0	0	0	400
S.5	7	140	0	0	108	60	85	0	400
S.6	30	143	134	48	45	0	0	0	400
S.7	26	131	145	41	57	0	0	0	400
All	81	687	492	294	623	334	143	146	2800
	(3)	(24)	(18)	(11)	(22)	(12)	(5)	(5)	(100)
Correct entries			492	294	623	334	143	146	2032
			(24)	(14)	(30)	(16)	(7)	(9)	(100)

- S. 1. ___ I leave now?
- S. 2. ___ I pour you a cup of tea?
- S. 3. ___ you lend me a hand?
- S. 4. ___ he help me ?
- S. 5. ___ I make a suggestion?
- S. 6 He said , "Do you expect that Mary ___ help?"
- S. 7 She said, " ___ he arrives at noon?"

Figures in parentheses are respective per cent shares.

Source: Field Survey was conducted during January-February, 1993 from a sample of 400 students of 10+2 level. These were sampled from seven colleges affiliated to Punjabi University: Three of these colleges are located at Patiala city, one college each is in the towns of Rajpura, Samana and Sangrur and one is located in the village Bardwal (Dhuri). For details refer Kaur (1993).

Table 2: Use frequency of entries in nine contextualized sentences by 400 students of 10+2 level from colleges affiliated to Punjabi University, Patiala, 1993.

Sentence location	No response	Incorrect entry	Entry of modal auxiliary		Total
			Appropriate	Exact	
1.	13	181	0	206	400
				[can]	
3.	47	177	0	176	400
				[can]	
5.	48	215	0	37	400
				[can/may]	
2.	40	299	19	42	400
			[may]	[should]	
7.	48	177	145	30	400
			[may/shall/can]	[should]	
6.	33	292	1	74	400
			[might]	[would]	
9.	34	223	104	39	400
			[will*]	[would]	
4.	36	192	79	93	400
			[would/won't/shall/can]	[will]	
8.	27	137	21	215	400
			[may/can]	[will/shall]	
All	326	1893	369	1012	3600
	(9)	(53)	(10)	(28)	(100)

C.1 How ___ I help you?

C.2 ___ I open a new box for you?

C.3 How ___ I do that with you, Sir?

C.4 ___ you take seventy paise?

C.5 You ___ return them all if they are sour.

C.6 But I thought you ___ buy at least two dozens.

C.7 Oh well, Sir ___ I pack twenty pieces?

C.8 I ___ come again tomorrow.

C.9 I said that I ___ take eighty paise.

Words in [] denote permissible modal auxiliaries in the context. Figures in () denote respective rounded up per cent shares.

* Its inclusion is partly aimed at giving the benefit of the absence of clause marker 'that' in Punjabi and partly to its significant use.

Source: Same as given in Table 1.

Table 3: Average correct performance rate of auxiliaries used in the non-contextualized and contextualized sentences by 400 students of 10 + 2 level from colleges affiliated to Punjabi University, 1993.

Modal auxiliary used	Sentences	Total number of Correct entries	total	per cent correct entries
----------------------------	-----------	------------------------------------	-------	--------------------------------

Non-contextualised set

Can	7	623	2800	22.25
Will	4	492	1600	30.75
May	3	334	1200	27.83
Would	4	294	1600	28.38
Shall	2	146	800	18.25
Should	3	143	1200	11.92

Contextualized set

Can	3	519	1200	43.25
Will*	2	308	800	38.50
Would	2	113	800	14.12
Should	2	72	800	9.00

* In one sentence 'shall' is equally admissible and hence its few entries are included.

Information is compiled from Tables 1 & 2.

Next, let us take the constructions with modal auxiliaries **will** and **shall**, the two most common ways of indicating likelihood and expressing futurity. Between the two, **will** (where admissible) is preferred to **shall**. For instance, in the non-contextualized set, the average rate of performance of **will** in the four sentences (S.3, S.4, S.6 & S.7) is 30.75 per cent and of **shall** in the two sentences (S.1 & S.2) is only 18.25 per cent (Table 3). In the contextualized set, average performance rate of **will** in the two admissible sentences (c.4 & c.8) is 38.50 per cent. A similar tendency (though in a different context) was noted by Quirk et al (1985: 136) when they observe that "**will** is used four times per thousand in spoken BrE" (British English) whereas "**shall** is used three times per ten thousand words in written English".

Apart from **will** and **can**, the other popular form of modal auxiliary is **may**. The average performance rate of **may** in the three non-contextualized sentences (S.1, S.2 & S.5) where it is admissible is 27.83 per cent (Table 3).

4. Concluding Remarks

To sum up, the frequency of individual modals varies greatly. **Will**, **can** and **may** are notably more frequently used by Punjabi ESL students than other modals like **shall**, **should** and **would**. It is noteworthy that **should** in the sense of polite request is replaced by **may** or **can** and **would** by **will** and **can**. Another noteworthy feature is in the context of the use of **will**, instead of its past tense form **would**, in the reported speech (C.9). As many as 104 students (26.00 per cent) used **will** against 39 students (9.75 per cent) who used the correct form **would**. Due to the absence of change of pronominal form and tense form in indirect speech in Punjabi, the learners of English do not seem to pay enough attention to these aspects of indirect speech in English.

Notes

This paper draws heavily on Chapter V of author's Ph.D. Thesis entitled *Aspects of Pedagogical Linguistics: A Case Study of English in Punjab*, submitted (November, 1993) to the department of Anthropological Linguistics and Punjabi Language, Punjabi University, Patiala.

Author is indebted to Professors V. Prakasam, Gurkirpal Singh Sekhon, Joginder S. Puar and S.S. Joshi for useful suggestions on earlier drafts of the work.

1. **Ought** is sometimes regarded as a modal rather than **ought to**. **Ought** is then said to be followed by a 'to-infinitive' (Collins:1990:219). The traditional category of the English modals is presented as eight in number (Palmer, 1965: 19). They include such monosyllabic words as: **may**, **can**, **shall**, **will**, **must**, **need**, **dare** and **ought**. Later, Brasnett, (1968-69:67) added four more modals to this list. These are **could**, **might**, **should** and **would**. **Dare** and **need** are semi-modals. They function as modals as well as full verbs.

2. Modals are used in negative structures to say that an action is forbidden because of a written rule, law or agreement or because it is unacceptable or undesirable. Negatives are formed by putting a negative word such as **not** immediately after the modal. After **could**, **might**, **must**, **ought**, **should** and **would** 'not' is often contracted to enclitic particle **-n't** and is added to the modal. **Can not** is usually as one word **cannot** (Collins:1990:219). **Shall**

not, will not, cannot, would not, could not, and shouldn't. May not is not shortened at all.

3. Modal auxiliaries as operators admit inversion, i.e. the subject noun phrase and the auxiliary (the first auxiliary if there are two or more) change place in a(n) (interrogative) sentence. Interrogative questions are, thus, formed by putting the modal in front of the subject (with exception to 'ought to'). Modals are used in tag questions. They are sometimes used on their own without a verb following them. For instance, in a tag question a modal which is used in the preceding sentence is repeated.

4. The term **code** refers to an auxiliary which picks up a full verb of a sentence referred to earlier (Palmer:1965:23). As a term meaning avoidance of repetition, 'code' was coined by J.R. Firth. Appropriate choice of Tense (modal, have and be + (n't) without verb) and pronoun in place of the subject NP in responses are the basic ingredients for the use of **code**.

5. Verb forms have different functions in finite and non-finite verb phrases. On this, the -s form and the past form are called **finite**, whereas the -ing participle and the -ed participle are called **non-finite** (Quirk et al., 1985:96).

6. British English prescriptive tradition forbids 'will' as a future auxiliary with 'I' or 'we' but this prescription is old-fashioned and is now widely ignored (Quirk et al., 1985:230).

7. Besides modal sub-category, the other sub-categories of operator in Punjabi are primary operator, passive operator and progressive operator (Puar:1990:29).

Bibliography

Brasnett, Clive. 1968-69. "Some Notes on the Modal Auxiliaries". *Bulletin of Central Institute of English*. No.7, 65-71.

Collins Cobuild English Grammar. 1990. Sinclair, J. et al. (Eds.) London: Hammersmith (Indus: Harper Collins Publishers India).

Hakutani, Y. 1972. "English Modal Auxiliaries: A Reconsideration". *Linguistics*. Vol. 90, 11-19.

Halliday, M.A.K. 1976. "Modality and Modulation in English". In G.R. Kress, (ed.) *Halliday: System and Function in Language*, London: Oxford University Press, 189-213.

Kaur, Setjit. 1993. *Aspects of Pedagogical Linguistics: A Case Study of English in Punjab*. Unpublished Ph.D. Thesis. Patiala: Department of Anthropological Linguistics and Punjabi Language, Punjabi University.

Leech, Geoffrey N. 1971. *Meaning and the English Verb*. London: Longman Group Ltd.

Palmer, F.R. 1965. *The English Verb*. London: Longman Group Ltd.

Puar, J.S. 1990. *The Panjabi Verb: Form and Function*. Patiala: Publication Bureau, Punjabi University.

Quirk, R.S. & Greenbaum, G. Leech & J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Note

**THE LANGUAGE OF ADMINISTRATION:
OFFICIALESE MALAYALAM - HINDRANCES IN USAGE**

B. SREEDEVI
University Of Calicut

India's national development since Independence has been a slow and steady process. National unity, an essential requirement for progress, arises from mutual co-operation, for which adequate communication plays a crucial role. A measure of the efficiency of the National planning system is the success with which mutual co-operation and understanding are built up. This is different from the meeting of basic needs and economic development. Language, like other cultural and anthropological factors, is a strong thread that ties people together. This does not mean that any one single language is the sole element or the most important factor that helps the unity of the nation. The use of the regional languages in administration to establish a true relationship between the people and the Government has been generally recognized to be a very important factor.

Governmental measures to develop the Regional Language in Kerala

Government of Kerala appointed a Committee in August 1957 to report on the measures to be taken for the progressive adoption of Malayalam as the official language of the State. The Committee submitted its Report in 1959; the Government have generally accepted its recommendations. In May 1969, Government appointed a Special Officer to attend to the work relating to the introduction of Malayalam as the official language. The proposals of the Special Officer are grouped here under the following four heads, viz., 1. General, 2. Specification of Departments where Malayalam may be introduced up to and including the District level, and in some measure, at the Head Offices, 3. Specification of Departments where Malayalam may be introduced up to and including the Taluk level, and 4. Ad hoc measures. These are discussed below.

1. General: These recommendations were intended to be applicable to all offices of the Government including the Departments of the Secretariat. They were to the effect that:

i. Communications issued in reply to communications received in Malayalam shall be in Malayalam;

ii. Interim communications involving the exercise of statutory powers shall be issued in Malayalam wherever the parties concerned have submitted their representations in Malayalam; final orders may be issued in English, but then, a gist of the final statutory orders shall be given in Malayalam to the parties;

iii. Notifications of general interest to the public shall be published in the Kerala Gazette in Malayalam;

iv. Minutes and records of discussions of all meetings of Advisory Committees, except at the State level, shall be kept in Malayalam.

2. Specified Departments (up to District Level): In the following Departments, with which the people have the maximum interaction, Malayalam shall be used upto and including the District level:

i. Panchayats

ii. Municipalities

iii. Co-operation

iv. Harijan Welfare

3. Up to Taluk Level: In the case of the following Departments, the use of Malayalam was to be enforced upto the Taluk level:

i. Land Revenue; ii. Excise & Prohibition; iii. Development; iv. Civil Supplies; v. Forest; vi. Weights & Measures.

4. Ad hoc Measures: The following measures were applicable generally for specific items of Government communication with the public:

i. At all levels, orders in pension cases shall be issued in Malayalam, and all correspondence with the applicants shall be in Malayalam;

ii. Forms used for the issue of permits and licences shall be progressively switched over to Malayalam;

iii. Application forms for "Learners' licences" (driving) and Drivers' licences shall be in bilingual (English/Malayalam) forms;

iv. All road signs, cautionary and statutory, shall have the necessary Malayalam write-up also;

v. All publicity programmes, except those of the Tourist Department, should be increasingly in Malayalam;

vi. Sign-boards on all roads announcing Town and village name, Block and Panchayat Limits, etc., shall also have write-ups in Malayalam;

vii. All offices in the State must feature name boards with Malayalam write up also.

Provision for use of Minority Languages. It was also laid down that similar instructions would apply to the use of the minority languages in the State, viz., Kannada and Tamil, wherever the speakers of these languages respectively accounted for 25% of the local population, or wherever such use of the minority languages was found necessary.

Implementation

In order to implement these measures, Government undertook the following steps:

i. "English" typists were trained in "Malayalam" typing also.

ii. Option was given to choose Malayalam at examinations conducted by the State Public Service Commission.

iii. Incentives were awarded to stenographers who acquired qualifications in Malayalam typing and shorthand.

iv. The Kerala Official Language (Legislation) Act 1969, amended in 1973, provided for the adoption of Malayalam and English as the languages to be used for the official purposes of the State of Kerala.

v. A Committee was constituted with the Chief Minister as its Chairman, for facilitating the use of Malayalam as the official language of the State.

vi. A Technical Committee was constituted to finalize the design and adoption of a typewriter key-board in Malayalam.

vii. Steps were taken to prepare the State Administration Report in Malayalam also.

viii. A Committee was constituted for the preparation of a "Malayalam Shorthand Manual".

ix. Steps were taken to have codes, manual, books, and publications of various Departments translated into Malayalam.

x. Malayalam typewriters were supplied in the ratio of 1 Malayalam to 4 English typewriters to all offices. Replacements of English typewriters were to be only with Malayalam machines.

xi. Orders were issued that display inscriptions on commemorative stones in public buildings, bridges, etc., should also be in Malayalam.

xii. Government directed that all Civil Courts and Criminal courts subordinate to the High Court of the State may use either English or Malayalam for writing judgements and other proceedings in such courts.

xiii. In consultation with the Director of Public Instructions, training in Malayalam shorthand writing was imparted through the State Institute of Languages.

xiv. Steps were taken through the State Institute of Languages to prepare necessary "forms" for all official purposes in Malayalam.

xv. The Official Language Committee of the State was reconstituted in 1973 with the Chief Minister as the Chairman.

xvi. A short guide book was published for the day-to-day reference in offices.

xvii. A Government notification was issued allowing the use of either English or Malayalam, at the discretion of the concerned official, for writing notes in the office files in the Secretariat Departments.

xviii. An English-Malayalam Glossary of Administrative terms was prepared by the State Institute of Languages for sale at nominal cost.

xix. A new Malayalam typewriter with keyboard using the "reformed" script was adopted for use.

Orders have been issued initially for 31 Departments, and later, for 15 more, stipulating the use of Malayalam. From *1 September, 1985* onwards, the Departments of (1) Panchayats, (2) Municipalities, (3) Forest, (4) Co-operation, (5) Harijan Welfare, (6) Tribal Welfare, (7) Fisheries, and (8) Archaeology had started using Malayalam. From *1 November, 1987* onwards, the Departments of (9) Jails, (10) Hindu Religious & Charitable Endowments, (11) Public Relations, (12) Weights & Measures, (13) Dairy Development, (14) Animal Husbandry, (15) Housing Board, (16) Social Welfare and (17) Registration, also started using Malayalam. From *1 November 1989* onwards, the Departments of (18) Revenue, (19) Agriculture, (20) Indigenous Medicine, (21) Water Transport, (22) Survey & Land Records, (23) Taxes, (24) Employment & Training, (25) Cultural Publications, (26) Museum & Zoos, used Malayalam.

A "Text Book"

In 1983, the State Government published a book with the intention of training the Government officials in the use of Malayalam for official purposes. The book was to be used as a reference in offices. The text included models of official notes, correspondence, and orders, typical of those in common use. "Correspondence" included the usual forms such as letters, circulars, orders, etc. A list of the Heads of the Departments was included. The book promised to make modifications as may be required in

the future. Arrangements were also made to devise various forms of application, etc., individually for the several Departments.

The following sampling of the Contents of the book shows how comprehensive it was intended to be. It covered the styles to be used while submitting files to higher officials; for making references to other departments and sections, the styles to be adopted by "Higher Officers" while recording their minutes, Drafts of communications, styles and usages to be adopted for disciplinary proceedings, letters including "Demi-official", Memoirs, Under-office notes and memoirs, circulars, teletypewriter messages, orders including "Office orders", notifications, notes regarding the conduct of meetings, programmes of visit of dignitaries, invitation letters, car passes, etc.

The State Institutes

The *State Institute of Encyclopaedic Publications* founded in 1961 has played an important role in making available knowledge accessible in Malayalam. Its main preoccupation is in two Projects: 1. The preparation of the most comprehensive and authoritative Encyclopaedia in Malayalam, in 20 Volumes, and 2. The preparation of an Encyclopaedia on World Literature, in 10 Volumes. It has already published NINE Volumes of the first project, the first five covering words starting with the vowels in Malayalam.

The *State Institute of Languages* was established in 1968 with the intention of developing Malayalam so as to make it a medium for the transmission of modern scientific knowledge. The Institute has mainly concentrated on the preparation of text books and other related materials for University students. The work began with the preparation of glossaries for Social Sciences, Biological Sciences and similar specialized areas. This enabled the students to exercise an option to write their examinations in Malayalam, up to the Pre-degree level by 1975 and to the Degree level by 1977. In addition, the Institute prepared programmes for the preparation of literature, enabling the setting up of several "corners" for Agriculture, Forestry, Co-operation, Law, etc., in the village libraries, as well as in higher echelon libraries. By 1986, it had published a thousand titles; in the next two years, it added another hundred, in subjects as diverse as Natural sciences, Physical Sciences, Engineering, Linguistics, Language, Literature, Culture, Arts, Social Sciences, etc. It regularly publishes a few journals also.

Administrative Glossary: Of special interest in this context is the *Administrative Glossary* published by it, covering 230 pages with 20 words to the page at an average. In addition, it has coined technical terms for more than two hundred thousand items in various subjects and disciplines.

The *Kendra Sahitya Academy* and the *Kerala Sahitya Academy* are two other learned bodies doing a "good job" for the development of Malayalam. By 1980, the latter had translated more than 200 Malayalam books into other languages, the majority being into English, Hindi and Tamil. The Academy has also done very good work in preparing books in Malayalam for Children, with more than 500 titles by 1984.

An Assessment of Progress

In spite of all these, the adoption of Malayalam for official purposes has been a slow process. A Survey was conducted using the Questionnaire appended. It was found that English remains the main language of official communication in most of the offices covered. Malayalam is generally used in the "spoken situation" whereas in writing, Malayalam is employed in *less than ONE* percent of official communications. The following Table is a summary of findings:

TABLE: USE OF ENGLISH AND/OR MALAYALAM IN VARIOUS OFFICES IN KERALA.

Office	Language used in:	
	writing (E = English; M = Malayalam)	speech
1. Kerala Public Service Commission Dist.Office Calicut	E & M	M
2. Integrated Child Development Scheme	E & M	E & M
3. Kerala State Housing Board	E & M	M
4. Kerala Institute of Research (ST/SC) (KIRTADS)	E & M	M
5. Govt., Homeopathic College	E & M (To sub.M)	M
6. Food Corporation of India	E	E,M, Tamil, Hindi
7. Centre for Research in Water etc. (CWRDM)	E	E & M
8. Calicut Development Authority	E & M	M
9. Standard Chartered Bank	E	M,E, Tamil, Kannada
10. Sales Tax Office	E	E
11. Nirmiti Kendra (Disagrees about use of M as offl.lang)	E	E & M
12. Electricity Board	E	E & M
13. Govt., Training College	E & M	E & M
14. Barat Overseas Bank	E	E & M

15. Inst. Human Resource Developm. in Electronics	E	M
16. Collectorate	E & M	E & M
17. Sub Registrar	E & M	M
18. High School	E & M	M
19. Telecommunication Office	E, Hindi, M	M
20. Co-op. Jt. Registrar's Office	E & M (M to sub)	M
21. Central Prison, Trivandrum	E & M	E & M
22. Secretariat	E & M	E & M
23. Electrical Inspectorate	E & M (E to superior offices)	E & M
24. Irrigation Circle Office, Trichur	E	E & M
25. Telecommunications	E, Hindi	E & M
26. State Lotteries	E & M	E & M

The latest *Evaluation Report* on the implementation of the official language policy has also shown that, except in a very limited number of Departments, English is predominant. The Chairman of the Implementation Committee has repeatedly emphasized the need to implement Malayalam. An effort was made to ascertain the reactions of the employees of the various offices on the issue. Their reactions are summarised below:

1. There is difficulty in finding out suitable Malayalam equivalents for most of the English usages. The available glossary is not of much help. Regional variation also is a problem which hinders intelligibility. Word-to-word translation is not suitable. It is better to keep to the familiar English words as such.
2. Letters, instructions, and orders from the Government and higher authorities are in English; hence the reply also is furnished in that language.
3. The Official Registers in Malayalam have not yet been developed. The use of English has become a habit, and hence easier. Sudden switch over will consume office time.
4. Some of the Departments eg., dealing with Law and Order, Finance and Audit, Arms and Ammunition, find English more suitable.
5. English is brief, effective and easy.
6. Lack of Malayalam typists and typewriters is another problem.

7. English is considered to be more prestigious.

8. "On the whole", there is no objection to the use of Malayalam, but conditions are not fully favourable at present.

These recorded opinions of the office workers is a pointer towards flaws in the whole programme. Office employees are usually degree holders; in our education system, there is a thrust towards and bias for English. As the medium of higher education, it has combined the objectives of language learning and language use in administration. Our administrative system is a legacy of the British and was predicated on the all-India use of English pervasively as the language of administration. On the other hand, Malayalam has no such traditional use in administration.

The teaching of the mother tongue in our Universities is mainly literature-oriented; as a result, students lack proficiency in its use in certain registers like the official. While English composition classes teach language use in letter-writing, comprehension, etc., Malayalam grammar classes deal with prosodies and features of creative literature. We use English for wedding invitations, leave applications, notices, etc. and feel more at home with it in matters of law courts and scientific knowledge. Thus, when trying to use the regional language for official purposes, there is an attempt to "translate" the style and terminology also from English, which results in failure. It is better to put the idea first in Malayalam and then verbalize it in appropriate style independently of the English "model".

Moreover, there is the fear of the geographical limitations of the regional language with its implied threat of restriction of movements in search of jobs. The overwhelming dominance of remittances from abroad in the economy of Kerala combined with the prestige value attached to "English-medium" schools has created a psychology of preference for the foreign language and neglect of the mother tongue from the earliest stages of schooling. A large number of prestigious and expensive "English-medium" schools have proliferated all over the State to cater to this fashion. An extreme case of this syndrome was the recent, much-publicised instance in which a student in an "English-medium" school was punished by having the hair on his head shorn for having spoken to his school-mates in Malayalam, even though it was outside the class room.

The present educational system robs the student of his mastery over any one language. Many of the students and office employees opt for English because they do not have an equal command in their mother tongue for formal and official situations. As a partial solution to the problem, there should be a change in the method and content of teaching

of the mother tongue in schools and colleges. More emphasis should be devoted to the syntactic and pragmatic aspects of language. The employees should be recruited to the offices only after a need based test in the language.

Acknowledgement

I express my gratitude to Sri. T. Madhava Menon for his valuable suggestions.

SILVER JUBILEE REVIEW

International Journal Of Dravidian Linguistics

Twenty-five years of a journal's existence is neither too long nor too short. It has reached adulthood and now it looks back to January 1972 and forward to December 1996 and beyond. The genesis and growth of the International Journal of Dravidian Linguistics (*ijdl*) from uncertainty to scholarly acceptance and consolidation for continuity, in all humility, is the story of struggle and faith in building a journal.

When *ijdl* was brought out in January 1972, following a resolution of the first All India Conference of Dravidian Linguists held at Thiruvananthapuram to publish a professional journal, the feelings of the Editorial Board were those of a mother delivering her first baby. The financial stringency and the limited capability of the press to print a professional journal at Thiruvananthapuram were two major issues which plagued half a dozen young men involved in its publication. The very first article in the first number was an Address to the Dravidian Linguists by Suniti Kumar Chatterji, the doyen of Indian Linguistics. T. Burrow, the co-author of the Dravidian Etymological Dictionary, William Bright, who later became the editor of *Language*, the Journal of the Linguistic Society of America, the well-known theoretician Kenneth L. Pike and his co-worker Kent Gorden, the Paninian Scholar J.D. Singh, the popular Dravidologist Kamil Zvelebil and the young theoretician Jackendoff and others had contributed to the first number.

Two notes, one on Brahui by Kamil-Al-Quadri of Pakistan who continues to contribute on Brahui and another by D.N.S. Bhat, and a review were included in that number.

The then-Vice-Chancellor of the Kerala University, Dr. George Jacob, wrote a letter of felicitation in which he expressed his confidence that the Journal would prove to be an effective medium for Dravidian Linguists for their research publications. The editor in his prefatory note sought the indulgence of the world of scholarship for a few shortcomings in the issue quoting a Malayalam proverb in which an imperfect sketch of an opera house drawn by children on sand was being looked at by a master architect with a smile. Letters welcoming the journal were received as soon

as the first issue was despatched. Kamil Zvelebil warned that many such ventures started in India had only a short span of life, which should not be the case with *ijdl*. *ijdl* took this advice earnestly and made arrangements for its continuous publication in January and June every year.

In recent years, no standard publication, on Dravidian Linguistics in particular and, on South Asian Linguistics in general, issued from India and from other parts of the world, has failed to cite the articles of *ijdl*. Cases of republication of its articles by the authors in their books were many. The readers from several parts of the globe have rated *ijdl* as the foremost professional journal issued from India and one of the important journals in the world of Linguistics. Exchanges were arranged with leading journals and indexing agencies. This significant growth within twenty five years has been possible because of the academic contribution of the scholarly world and the dedicated work of half a dozen men for the production of the journal.

COVERAGE

Upto 1994, in 23 Volumes, *ijdl* has published 402 full-length articles, 118 notes, 6 reports of conferences, 103 reviews of books, 23 Presidential addresses of Conferences and 6 survey reports. Seminar reports and miscellaneous information have also been published.

In the early volumes, the practice of discussing the articles published in the previous issues was implemented in which many mature linguists including C.F. Hockett participated. This could not be continued because many comments were received late and the replies were delayed. In a modified form, the critical reviews of articles were published though not regularly. The discussions secured many citations in several other publications. It was recommended to be continued though reluctance was perceivable, especially among the authors of the main articles.

Classical articles which are rare and out of print were republished in the journal, with adequate acknowledgement to the source.

REGULAR READERS OF IJDL

The life members of DLA, who are regular readers, number 580, out of which 44 are from outside India. Among them, 22 are from USA, 3 from Sri Lanka, 2 each from Japan, France, West Germany, U.K., Netherlands and one each from Bangladesh, Canada, Malaysia, Italy,

USSR, Belgium, Switzerland, Finland and Australia. Besides, each issue of *ijdl* is sent to about 100 libraries and other research institutions in India and abroad. The annual members from India and other countries are not listed here.

LANGUAGES ANALYSED

The sister Journals in India had very little space for Dravidian Languages compared to classical and Indo-Aryan languages. The thrust of *ijdl* was to give preference to Dravidian languages, major and minor including tribal languages. Wherever possible it accommodated languages of other families.

Besides articles on the Dravidian languages, [Tamil (53 articles), Malayalam (35), Kannada (15), Telugu (9) and Tulu (13)], articles on Indo-Aryan languages including Sanskrit (18), Vedic Sanskrit (1), Prakrit (1), Hindi (9), Hindi-Urdu (1), Sindhi (1), Kashmiri (3), Urdu (2), Saurashtri (1), Marathi (2), Konkani (4), Bengali (5), Hindusthani (1), Oriya (1), Central Pahari (1) and Punjabi (1) were also published.

Among the tribal languages in India, articles on Irula (1), Jenu Kurumba (1), Aalu Kurumba (1), Cholanaickan (1), Muduga (1), Paniya (1), Adiya (1), Malto (1), Brahui (2), Kurux (1), Gondi (1), Kubi (1), Dhangar or Jhangar in Nepal (2), Tibeto-Burman (1), Apatani (1), Tibeto-Himalayan (1), Mahto (1), Kinnauri (1), Remo (Bonda) (1), Kharia (1), Khasi (1), Bhili (1), Gojri (1) and some Andamanese languages were published.

Among foreign languages, articles on English (9), American English (1), Chinese (1), Sinhalese (2), Creole Portuguese (1), Lithuanian (Baltic) (1), West Greenlandic (1), Bangala (1), Arabic (1), Japanese (1), Russian (1), Fijian (1) and Magyar (1) were accommodated.

Dravidian-Indo-European (2), Dravidian-Negro-African (1), Indo-European (4), Indo-Aryan-Tibeto-Burman (1), Elamite-Dravidian (1), Dravidian-Sinhalese (1), Dravidian-Munda (1), Indo-African and Indo-Iranian (1) and South Asian (1) were reported.

Apart from these, 71 general articles, touching all aspects of linguistic theory without specifying any particular language, have also been published in *ijdl*. The Bibliography, prepared and published in Vol.XXII (No.1) by K. Balachandran, gives the captions of the articles with the index

of authors. Unlike other journals in India, theoretical articles, reflecting all shades of opinions including Transformational Grammars and their derivatives, have been encouraged. It was Prof. Southworth who expressed a desire that *ijdl* should be a bridge between the Russian Theories of Linguistics and those of USA and Great Britain when the cold war was at a high pitch.

Besides well-known contributors in the discipline from all over the globe, about 20 per cent of the authors were very young researchers who had contributed for the first time to the field of linguistics in *ijdl*. This is a matter of satisfaction for the Editorial Board.

NATIONALITY OF THE CONTRIBUTORS

In *ijdl* Vol. I to XXIII, from 1972 to 1994, out of the 402 articles, 14 articles have co-authors. Among the authors, 265 are from India, 65 from USA, 12 from Sri Lanka, 8 from Netherlands, 7 each from Australia and U.K., 5 from USSR, 4 from Germany, 3 each from Japan and Czechoslovakia, 2 each from Malaysia, Nepal, Pakistan, Singapore, Denmark, Austria, one each from France, Iraq, Senegal, Sweden, Canada and Italy and 4 from other countries. With regard to Book Reviews, a total of 103 books have been reviewed. Among them, 18 are foreign publications. They have been reviewed by 96 Indian scholars and 7 foreign scholars.

CITATIONS OF THE JOURNAL IN OTHER PUBLICATIONS

Journals and books on Dravidian, and on Indian Linguistics and Linguistic Theory, had cited articles of *ijdl* or republished them with permission. Important among them were the International Journal of Sociology of Language (Vol.XVI), Mouton (Netherlands), Studies in Linguistic Sciences (Vol. VIII, No.1), Indian Linguistics (Pune) Language (USA), Journal of American Oriental Society, in the series on Syntax and Semantics, had briefed two articles and published them. The articles of William Bright, Colin Masia and M.B. Emeneau were republished in book form. James W. Gair and S. Suseendra Rajah have also requested permission to reproduce their articles. The books, *Pioneers in Linguistic Series - I*, *Sri Lanka Creole Portuguese Phonology*, *Effectivity and Causativity in Tamil*, *Language and Brain* and *Dynamics of Verbal Extension in Tamil* are collections of articles of *ijdl*. The Linguistic Abstracts had summarised articles from *ijdl* with bibliographical index. Specifically, no book on Dravidian published within two decades had failed to quote *ijdl*.

IMPORTANT SUBJECTS DISCUSSED IN *ijdl*

Though a detailed analysis of the articles will throw more light on the theoretical and applicational aspects, we are compelled to list only the titles due to want of space. Two monographs containing the important articles, already published, are to be compiled, in due course, on the basis of theory and application, to give an overall view of the thrust of the articles. This, we propose to do within a short time.

Tamil

Relevance of the terms *mey*, *orru* and *pulli*, to the systems of Tamil Morphophonemics, Oral vowels of colloquial Tamil, Stops and Gemination *ai* and *au*, the Palatal Nasal, the *aḷapeṭai* forms, words denoting colours in *Cilappatikāram*, Adjectives in Modern Tamil, Manner adverbs, Verbals, Nominals, Effectivity and Causativity, Causativity and Verbal base, Negation, *Eluttu*, *pull* and *mey* in Tamil Grammar, Jaffna Tamil Verbal System, Dynamics of verbal extension in Tamil, Homogeneity and integrity of *Tolkāppiyam* and *Pāṇiniya sikṣa*, Dynamics of Kampan's language, Locative terms in Tamil, Vowel shift in Tamil Diglossia etc., have been dealt with.

Malayalam

Enunciative vowel, two Malayalam Phonemes (*n*) and *n*, Nasal assimilation, Phonological systems, Present tense markers, Modern Malayalam tense form, Phonological formulae for verbal suffixes, Aspectual system, Pronoun and relative clause in Laccadive Malayalam, Two Malayalam inscriptions from Coorg, Malayalam grammar by Carmelites, Degrees of politeness in Malayalam, Early use of the word Malayalam, Negative finite verb in Malayalam, classification of Malayalam verbs in the Malayalam Lexicon, Auxiliary verbs in Malayalam, *Deictic* properties of *person* in Malayalee ESL learners, Ellipsis in Malayalam, Minority language situation in Kerala, the Personal pronouns and pronominal forms in Malayalam, etc., are discussed.

Kannada

Must and *Vectors*, Palatalizing and Velarizing dialect, Nasalization, Compound verb construction, Kannada Syntax, *Pala-Vela* dialect of Kannada, Palatalization and Velarization in Kannada dialect, Kannada Grammars and Sanskrit Tradition, Influence of Sanskrit on Kannada

Compound formation, Agent suppression in Kannada and Tulu, Lexical opposites in Kannada, Bilingualism in a Karnataka border village, Woman and language use etc., have been covered.

Telugu

Intonation of colloquial Telugu, Secondary verb in Telugu, Vowel harmony in Telugu, Status of aspirated sounds and their synchronic and diachronic changes, Nominal inflection in the Sri Lanka Gypsy Telugu, Names and address terms in Telugu, Tense shift in Salem Telugu due to contact with Tamil, Development of Telugu lexicography, English loan words available in Telugu etc., are the most important articles that have been published.

Tulu

Negation in Tulu, Auxiliary verb in Bunt's Tulu, Nominative in Tulu, Agent suppression in Kannada and Tulu, Sociolinguistics and Tulu dialect survey, Folk culture reflected in Tulu place names, Linguistic studies on Tulu grammar etc., have found a place in *ijdl*.

COMPARATIVE DRAVIDIAN

Enunciative vowel, Evolution of dental nasal, PDr. Morphophonemics, Quantitative alternations in Dravidian, Numerals in Dravidian pronominal suffixes, Verbals, Gender and Number, PDr. word for horse, Intent constructions, Sub-groups in Dravidian, Close-relationship among Dravidian and the descent of Dravidian, Inter-relation between Tulu and other Dravidian languages etc., are among the significant articles.

DRAVIDIAN TRIBAL LANGUAGES

Articles include aspects in *Kurux*, PDr. \bar{O}^* and Malto, General idea about Brahui, Gondi aspectual system, Demonstratives in *Kūbi*, *Kūvi* and *Kūi*, *Emic* relation between non-transitivity clauses, Paired sentences reversals in Dhangar (*Kudux*) of Nepal, Echo word formation among the Aalu Kurumba, Tenseless verb in *Jenu Kurumba*, *Inūla* place names, Language of Cholanaickan, Language and culture of Mudugas, Personal Pronouns in the *Paniya*, *Adiya* speech forms of Kerala etc.

RELATION BETWEEN DRAVIDIAN AND OTHER GROUPS

Dravidian and Indo-European: the neglected hypothesis, Dative case in Indo-Aryan and Dravidian, The usage of Sanskrit grammatical

Terminology employed to describe Tamil, A study on the relationship between Tamil and Japanese: Intervocalic stops in two languages, Tamil sound sequence of the Japanese alphabet, Comparative study of Japanese and Tamil vowels, Influence of Dravidian Phonology in Sinhalese, Connection between Greek and Dravidian words and Indo-European and Dravidian words, Dravidian and Negro-African, Dravidian and Munda etc., are some of the topics covered.

Sanskrit

Verbal stems and syntact structures in Sanskrit, Grammatical methods in *aṣṭādhyāyī* and *Tolkāppiyam*, Linguistic speculations of Sanskrit rhetoricians, Definition of sentence: Bharthruhari's view, the treatment of *Kāraka* by *Pāṇini*, the influence of Sanskrit on Kannada compound formation, Sanskrit influence on Tamil grammar, *Tolkāppiya pirappiyal* and *Pāṇiniya śikṣa*, The emergence of Sanskrit up to the era of Modern Indian Languages, Treatment of lexical relations in the Indian traditions, Thomas Burrow's contribution to the study of Sanskrit, *Ṛgvedic* loan words, Exceptions and synchronic analogy in sanskrit, Historical change and synchronic structure in the case of Sanskrit vocative singular *ā* stem and root nouns, *Pāṇini's theory of language and Kāraḥas*, Technical terms in *Pāṇini*, Notes on Jaimini *Ṛgveda*, Phonological change in *Pāṇini*: his theoretical assumptions etc., have been featured.

Hindi, Bengali, Pahari, Punjabi, Gujarathi, Urdu, Konkani, Kashmiri, Hindustani, Oriya, Saurashtri, Sinhalese

Consonant assimilation in Hindi, Pre-lexical nature of the causative rule in Hindi, Characteristics of intensifiers in Hindi, Reflexivization in Hindi, Use of particles with different shades of meaning, Pidgnization in Hindi-Urdu, The conjunctive participle in Hindi-Urdu, Mantissa of indefinie words in Bengali, Bengali verb, Reflexivization in Bengali, Bengali syntax, Flirting as an asymmetric speech act in Bengali, Bengali syntax, Time factors in trade talk, Non-Aryan elements in a Pahari, Linguistic atlas of Punjabi, Casual purposive predicates in Gujarathi, Palatilization in Konkani, Shennoi Goembab's Konkani grammar, Konkani - a case for language planning, Model verb of obligation in Kashmiri, Compound verb in Kashmiri, Problems of MT maintenance in Multilingual setting: the Kashmiri case, Two Koines compared: Guyanese Bhojpuri and Calcutta Bazaar Hindustani, A study of Englishization of Oriya, Dravidian interference in the Saurashtri secondary verb, The verb in Sinhala,

Convergence and diglossia in Sinhalese pronouns etc., are among the topics dealt with.

Foreign languages

Creole Portuguese phonology, English in Indian situation, Features of Black English as a representative of a non-dominant culture, Indian loan words of medieval period English; A socio-functional approach to the teaching of English in India, Error analysis of wrong usage of English noun phrases, What's American about American English, Consonants in West Greenlandic. Lithuanian Baltic *ē-* or *a-* preterite, Morphosyntactic features in South Asian languages, Convergence in South Asian languages, Description of Sun or Moon in Arabic, Implications of ancient Chinese retroflex endings, Causal factors in the loss/maintenance of minority language: the case of Tamil in Fiji etc., have been included.

Indo-Aryan and Indo-European

Compound verb in North Indian languages, Linguistic stratigraphy of North India, Indo-European verb morphology, Old and middle Indo-Aryan verb, Verb in Indo-European Syntax, Indo-European Tibeto-Burman connection etc., are some of the articles.

General-Theoretical

Apart from the above topics, articles on linguistic problems in general without specifying any particular language were accommodated in *ijdl*. Some of them are Tamil-Saurashtri bilingual grammar, Tamil-Malayalam functional aspects, Learners uncommunicativeness, caste and languages, Sociolinguistic stratification and speech variation, Linguistic convergence, Subgrouping, Defining standard language, Standardization process, Structural anthropology, Generative semantic methods, Semantics, Systemic phonology, Structural folkloristic and morphological analysis of Oroan narratives, Lexical insertion, Concept formation and language equivalences, Inductive area-genesis, Autonomy in language, Process grammar: an integrated theory of acts, speech acts and language structure, Abbreviations and acronyms in Indian setting, Philosophical musings of our ancient linguists, Dialect studies planning for the future, Focus and functions of applied linguistics, Myth of Malleeswara Peak, Computer and translation, Nature of grammatical competence etc.

FINANCIAL POSITION

To begin with, from 1972 to 1975, the University of Kerala, through the Department of Linguistics, provided funds for the publication of *ijdl*.

The sale proceeds were remitted to the University. The income and expenditure were so widely variant that the auditor appointed had to point out the difference to the University. In 1976, taking note of the financial difficulties and the delay in the publication of *ijdl*, the ICSSR, New Delhi, on our request, came forward to create an endowment of a lakh of rupees, by contributing Rs. 25,000/- from its funds, Rs. 50,000/- from the Ford Foundation routed through the ICSSR and Rs. 25,000/- contributed by the Dravidian Linguistics Association. A condition, that 75 per cent of the life membership fee after the creation of the endowment should be remitted to the endowment, was accepted. The interest that accrued from the endowment, which now stands at one lakh and forty thousand rupees, was just enough in the 1970s for printing and publishing the Journal. The endowment, created mainly due to the interest of Dr. J.P. Naik, the then-member-secretary of the ICSSR, gave sufficient independence to the Journal. Even today, financially, *ijdl* has not overcome the difficulties, because of the high cost of paper and printing. Each number costs Rs. 20,000/-. The postage will come to Rs. 3000/- for each number. Only a nominal sum is spent for the payment of the staff for clerical work. But the hope that it can overcome the financial stringency soon, is very pronounced in the minds of the members of the Editorial staff.

EDITORIAL BOARD

An International galaxy of distinguished scholars, S.M. Katre (Pune), T.P. Meenakshisundaram (Coimbatore), P.B. Pandit (Delhi), H. Berger (Heidelberg), R.C. Hiremath (Dharwar), Leigh Lisker (Pennsylvania), K. Mahadeva Sastri (Anantapur), Bh. Krishnamurthi (Hyderabad), J.R. Francisco (Philippines) T. Kandiah (Ceylon), John Marr (London) and Semyon Rudin (Leningrad), adorned the first Editorial Board. Later, in the place of Semyon Rudin and P.B. Pandit, M.S. Andronov (Moscow) and R.N. Srivastava (New Delhi) were nominated to the Board in 1974 and in 1976 respectively. The Editorial Board was reconstituted in 1981 with new members, namely Sebastian Shaumyan (USA), E. Annamalai (Mysore), Noboru Karashima (Japan), Harold Schiffman (USA), H.S. Anantanarayana (Osmania), F. Gross (France), K.V. Subba Rao (Delhi), D.N.S. Bhat (Thiruvananthapuram) and J. Vachek (Czechoslovakia). The same board continued for ten years and in June 1991, the Editorial Board was again reconstituted with James W. Gair (USA), John F. Marr (U.K.), A. Govindankutty (Netherlands), N. Gurov (USSR), M. Israel (Madurai), G. Sambasiva Rao (New Delhi), B.P. Mahapatra (Calcutta), K. Rangan (Tanjore) and E. Annamalai (Mysore). In 1994, K. Kapp (Germany) was also included in the Editorial Board.

THE ENDOWMENT COMMITTEE

To oversee the endowment, a committee, with K.N. Raj (1972-86) and Hans Raj (1987 onwards), nominated by the ICSSR as its representative, and R.C. Hiremath (1972-86) and G.K. Panickar (1987 onwards), by the Dravidian Linguistics Association as another member, and the Editor of *ijdl* being the Member Secretary, was constituted.

EDITOR AND HIS ASSOCIATES

The Chief Editor, V.I. Subramoniam (1972-90), was assisted by the Deputy Editors, V.R. Prabodhachandran Nayar (1972-73), Elias Valentine (1973-78), P. Ramachandran Pillai (1974-80), P. Somasekharan Nair (1974-80), S. Velayudhan (1981-88), V. Prakasam (1981-90), E. James R. Daniel (from 1981), B. Anvita Abi, B. Sreedevi and Annie Monsy (from 1991), and V. Syamala (from 1993). A.P. Andrewskutty, Deputy Editor from 1978, became Associate Chief Editor in 1987.

V.I. Subramoniam became the Hon. Editor when A.P. Andrewskutty took charge as Chief Editor in 1991. They all looked after the subscription, bookings, payments, reviews and articles.

From 1972 till 1988, the Kerala University Co-operative Press identified itself with the production of the Journal. It has its limitations. Compositors were not trained in linguistics. Spelling mistakes persisted in spite of careful proof-reading. Later, computerised printing was adopted from June 1991, and the production is now of reasonably high standard and mistakes are few. It is our goal that each number should be mistake-free.

DECENNIAL CELEBRATION IN 1982

In the Gujarathi Samaj Bhavan at Delhi, the Decennial celebration of *ijdl* was celebrated when the 10th Annual Conference of Dravidian Linguists was convened by the Department of Linguistics of the Delhi University. Hon. P.V. Narasimha Rao, the then-Minister for External Affairs and a reputed scholar, under the Chairmanship of Hon. Thiru Lakshmanan, Deputy Speaker of the Lok Sabha, inaugurated the Conference. Dr. Rao, in a thrilling address, praised the constructive activities of the Dravidian Linguistics Association. Dr. B. Ramachandra Rao, the then-Vice-Chairman, University Grants Commission, felicitated the delegates.

A stimulating decennial address by Prof. Sebastian Shaumyan of Yale University was heard by a full house. Dr. Karan Singh presided over the meeting and concluded that in spite of the foreign exchange difficulties experienced then by the Government of India, Prof. Shaumyan's air fare and a grant of Rs.10,000/- for a single lecture were met by the Dravidian Linguistics Association. K.V. Subba Rao, now Professor of Linguistics in the Delhi University, spared no efforts in making all arrangements for the Conference.

SISTER INSTITUTIONS

The All India Dravidian Linguistics Association which is responsible for projecting the Journal always extends a helping hand for its growth. Its life members receive a free copy of the journal.

Life membership fee is Rs. 750/-. For countries other than India, it is U.S.\$ 150/-. Institutional membership fee is Rs. 5,000/- or U.S.\$ 500/-. Annual membership is Rs. 200/- or U.S.\$ 40/- per year.

The International School of Dravidian Linguistics, another sister institution, provides the academic support for the journal. Its well-stocked library and its researchers help in evaluation of articles and check the references in this library, whenever a need arises.

REGISTRATION OF *ijdl* AS A SEPARATE BODY

Until 1973, *ijdl* was issued as an arm of the Dravidian Linguistics Association. It was registered as a separate society on 16.2.1973 to give an independent identity.

DLA NEWS

To report the activities of the Dravidian Linguistics Association and other institutions in India and abroad, a monthly, *DLA NEWS*, began publication from August 1977, and a copy each is sent to all life members and important organizations in India. To life members outside India, the twelve numbers are sent in a bulk due to high cost of posting. In India, each number costs, for posting, 15 paise due to postal concession whereas it costs Rs. 2/- per number if it is to be posted by Sea Mail to countries other than India.

FUTURE EXPANSION

In the Silver Jubilee year, *ijdl* has proposed to increase the pages to accommodate a greater number of research papers. The price of the

Journal, which now stands at Rs. 200/- or U.S.\$ 40/- per year, has to be revised to offset the high cost of publishing. Unless the corpus of the endowment is increased either through contribution or through increasing the number of life memberships, the rise in the price of each number is inevitable.

The Journal has a building of its own in the heart of the Thiruvananthapuram city. The total office space is 103 sq.m., built on 3 floors, on a piece of land admeasuring two cents gifted free of cost by the Dravidian Linguistics Association. The cost of construction, of about Rs. 2.5 lakhs, still remains unrepaid.

Countries like Australia, Malaysia, Singapore and South Africa are having very few life members. These and other countries have to be contacted for increased readership.

ijdl will also continue to serve as a window for all theories in Linguistics and for publishing contributions from all nations without any discrimination.

The Silver Jubilee Seminar, to be convened in 1996, will finalize the future policies and programmes for expansion of *ijdl*.

SILVER JUBILEE CELEBRATION

During the Silver Jubilee celebration, two Editors of leading journals of Linguistics in the world, will be invited to deliver felicitation address on that day. The total estimate for future expansion is approximately 10 lakhs of rupees. The scholarly world and the funding bodies in India and elsewhere, the Governments at the Centre and the States in India, it is hoped, will liberally contribute to the expansion and regular publication of *ijdl* which has become an institution in India for Linguistics in general and Dravidian Linguistics in particular.

Review

AN ANTHROPOLOGICAL ATLAS: *People of India, National Series Volume XI, by K.S. Singh, Published for the Anthropological Survey of India by the Oxford University Press, Delhi, etc., 1993; pp 156 + xii, price Rs. 700/-*

This much awaited volume of maps on the "People of India" (POI) will be welcomed by scholars of several disciplines. It is based on the primary data collected during the monumental project, and will be used as authoritative source of data for a long time to come. There are 72 maps on ecology-related and cultural traits, 16 on languages and linguistic traits and 61 on demographic and biological traits. Though some of these details had been shown in earlier maps released by the Anthropological Survey of India, the data now published are an "up-date" of these.

The foundation of the POI is the "community", a concept which has not been precisely defined anywhere. In fact the First Volume has discussed this concept and left it deliberately to form itself according to the perception of the segment of the people concerned (Singh 1992:23-29). This Atlas is also based on the "community". A trait if found in a "community" finds a place on the concerned map, irrespective of its numerical strength. The bar charts in the Atlas indicate the number of the communities showing a particular trait, and the length of the bars the total.

The "Introduction" is not exactly an example of clarity of expression. For example, the sentences: "...Thirdly, where the value of a community is less than fifteen, a height of 1 mm has been shown. The bars in the biological charts show the range of frequency of the given trait. Secondly (*sic*), language has been included as a category in the biological maps" (p 2), are unclear. The first of these sentences may be interpreted that, if the number of communities showing (or not showing) a trait is less than fifteen, then a space of 1 mm is allotted to such cases in the concerned bar chart. In the "second" sentence, which incidentally comes *after* the "thirdly", the statement that language is a "category" of biological traits seems to be "odd"!

While fully appreciating the great work that has gone into this publication, there are a number of cases, most unfortunately, where reasonable doubt is attracted on its credibility. A few are given below:

1. In Map 3, "Terrain" has been classified as Hilly, Plateau, Plains, etc. The basis has not been indicated; that it is probably arbitrary is indicated by the

fact that while some communities are shown as inhabiting "high altitude" areas in Maharashtra, Gujerat, Rajasthan and Andhra Pradesh, where the maximum elevation of the highest peak would be below 2000 meters + MSL, no community is shown as living at such altitudes in Kerala, where the "High Ranges" at above that elevation are the home of several communities. Similar remarks apply to classifications of climate (Map 4), humidity (Map 5), rainfall, etc. It would have been helpful had the basis of classification, (eg., High altitude = land above 2000 meters + MSL), been furnished.

2. In Maps 20 (consumption of cereals) and 21 (consumption of pulses), some of the terms, eg., *nagli*, *kulthi*, *moth*, etc., do not convey sense to a person not conversant with the vernacular in which these terms are used. The common terms, or the botanical nomenclature, would have been informative. In Map 23, showing the consumption of roots and tubers, the bar charts indicate that a few Scheduled Tribes, and "other" communities, probably less than fifteen, are reported to have "no consumption" of roots and tubers, *in Kerala*. This is hard to believe.

3. In the note under Map 34, showing polyandry, the mention that non-fraternal polyandrous marriages are reported from two communities in Kerala is not reflected in the map. It is presumed that the single community having fraternal polyandry shown against Tamil Nadu refers to the Toda; if so, the reportage must have been diachronic because according to Walker (1986:75), the practice has been long given up among them. If the map is intended to project diachronic information, then it should have mentioned several other communities all over India where polyandry was prevalent in the near or distant past.

4. It is interesting to note that marriage of ego with mother's brother's daughter/father's sister's daughter is practised among more than half the number of communities in India, and that there is no region in the country where it is totally absent (Maps 36 & 37). This may require the revision of many theories on the subject in current literature.

5. In the note under Map 59, Animal Husbandry has been shown to be the "current occupation" of 999 communities. It is not clear whether it is their "main" or only occupation. If not, the number of communities who engage in Animal Husbandry along with other occupations must be much greater. Perhaps this is an example of "overlap" which is unavoidable when such a lot of data has to be exhibited; such overlaps may be noticed in the maps relating to other traits also, for example in those relating to marriage, bride price and dowry, family structure, etc.

6. An example of how the information based on "community" could give a different picture from Census data is furnished in Map 66, showing

engagement in organized sectors. If at least one member of a community has such employment, then that community is counted into the concerned bar chart. Thus, in Kerala for example, less than 15 communities each among the ST, SC and "others" are unrepresented in this category. This of course has no bearing on the magnitudes of unemployment with reference to the *population*. The adoption of "community" as the base leads to overlap in the economic activity maps, generally, as at least some members of any community are likely to be found in each category of activity or occupation.

7. Again hard to believe is the revelation in Map 71 that *only* about 15 communities among the ST and SC in Kerala use "Modern Medicare" and that about half of the "other" communities use "Only Traditional" systems!

The Language Maps: While the above lapses may be excused, there is actual "dys- information" in regard to the linguistic maps. The worst case is Map 72; Kurukh (Index No. 39) and Kolami (No. 52) are exhibited as Austro-Asiatic languages, Kui (No. 41) and Malto (33) are marked over with vertical lines showing that they are in the Indo-Aryan family. Speakers of Gondi, Pengo, Maria, Gadaba, Konda, Parji, and Kuvi are shown as limited to Andhra Pradesh only. No community is shown as speaking any Dravidian language in Madhya Pradesh, Maharashtra, Orissa, Bihar or West Bengal. This map is not merely factually wrong, but also is totally misleading. The same bias is reflected in the note under Map 73. Care is taken to point out that "the presence of Dravidian languages reported in Rajasthan, Gujarat and Delhi is due to the *recent* immigration of communities speaking Dravidian languages" (*emphasis* supplied). The matching indication supplied to explain the presence of Indo-European languages in the Southern and North Eastern States mentions traders and nomads, pointing out that some of them are "historical" immigrants. If diachronic information was intended to be conveyed as an explanation to synchronic illustration in a Map, then the fact that several communities, eg., Bhils are likely to have been "originally" Dravidian speakers, whose language was later replaced by Indo-Aryan speech forms, and the dynamics of language change, eg., the emergence of pidgins like Halbi and Sadari, etc., should have been mentioned. It would have been best to have had avoided diachronic commentary in this Atlas.

In the note under Map 74, Hindi is said to be the "national language"; however laudable the sentiment, the statement is factually unsupported as the Constitution of India prescribes Hindi as only one of the two *official* languages of the Union. The inadvisability of adding comments to the maps is also illustrated by the confused narration under Map 75, showing bilingualism. The map clearly brings out the fact that more than half the number of communities in India, and less than half only

in Tamil Nadu, West Bengal, and Pondicherry, are bilingual. The explanation: "the monolingual communities are fewer ... in all the states of north-east India ... (because of) the larger number of tribal languages spoken in this region" requires further comment. The statement that the number of monolingual communities in Bihar, etc., is *less* because of "linguistic localization of the communities" is puzzling.

In Map No 76, the Perso-Arabic script is shown as being in use in most of the States *excluding*, among others, Lakshadweep where the overwhelming majority are Muslims. The use of Devanagiri is conspicuous by its absence in the Southern States - strange, if the data in Map 73 showing the ubiquitous spread of Indo-Aryan languages throughout South India are to be believed. Common sense indicates that there are at least a few communities in Kerala who use the Tamil script and some in Tamil Nadu who use the Malayalam script (examples: Tamil Brahmins and Kammalar in Kerala, and Kaniyar and Tarakar in Tamil Nadu respectively, cf Singh & Manoharan 1993:315,337,340). But this is not indicated in the map.

The Biological Maps: Valuable information has been furnished in the "biological" maps, even though here again, the basis of "community" may be confused with Census data in a few parameters, eg., Sex Ratio (Map 88). Map 89 (Child-Women Ratio) is probably on the basis of the Census data rather than of "communities". Maps 93 and 93A bring out the "remarkable" uniformity of distribution of stature among the populations of the western, central and southern states. Maps 93B and 93C bring out "the close correspondence between stature and social status in India".

It is remarkable that Rajasthan and the northern States have high proportions of Hyper Dolichocephalic and Dolicocephalic head forms; brachycephaly is marked in a south eastern sweep through Gujerat, Maharashtra, Karnataka and Tamil Nadu, leaving Kerala and Andhra Pradesh predominantly dolicho- and meso-cephalic (Map 94A). The earlier theories which associated "Dravidians" with the "long headed Mediterraneans" and "Indids" from brachycephals may have to be regarded as purely mythical in the light of this evidence. The tribal populations of Peninsular India are predominantly long headed, except in Karnataka, where they are mesocephalic (Map 94C).

The use of "language as a category of biological traits", as mentioned in the Introduction, is perhaps intended to cover Maps 95 (series) to 115 (series). Maps 95 (series) show that there is a progressive increase in the relative breadth of the nose as one proceeds southwards and that Scheduled communities tend to have broader noses. Irrespective of language, tribal communities tend to show a gradation from "medium" noses to broad from west to east, the tribes of Central and South India,

and of the Andaman and Nicobar Islands being mesorrhine ("medium"). In regard to Facial Index, inter-group differences within the same linguistic group seem to be more pronounced than between linguistic groups. This seems to be, generally, true of other morphological characteristics also. No clear correlation between language and the incidence of the non-taster gene (t) has been brought out, but it is comparative greater in Dravidian tribal communities, especially those in the Peninsula.

The data relating to the incidence of the Sick Cell trait (Hb.S) is interesting. It is very low or virtually non-existent among communities speaking Austro- Asiatic languages, and in the Himalayan belt. It is quite high among "Other communities" in Uttar Pradesh and Madhya Pradesh, while there is no spectacular difference in the rest of the States - it seems to be irrespective of whether the language is Indo-Aryan or Dravidian. It is absent among the Scheduled Castes in Kerala and Karnataka, but noticed in SC communities in Tamil Nadu and Andhra Pradesh. (There may be an error in the editing of Map 101, because a triangle signifying zero incidence, dot-coded for SC-Indo-Aryan, is shown against Kerala). Relatively high values are found for tribal communities, *irrespective of language*, in Madhya Pradesh, Gujarat, Karnataka and Kerala (Map 101A).

The Haemoglobin - E is limited to the North-eastern states, with a low incidence in West Bengal and Sikkim (Maps 102 & 102A), Indo-Aryan, Tibeto-Burman and Austro-Asiatic communities in this area exhibiting it. Dravidian communities in West Bengal however do not have it. Gene A under the ABO system is very high at 50.62 % among the Onge, and at 31.7% among the Sherpa (Sikkim); it is relatively high in the Himalayan belt and low over most of the rest of India. Pockets of high incidence are found among the Scheduled Tribes of Dadra & Nagar Haveli, Karnataka and Kerala, where the distribution of this gene among the tribes is higher than among Scheduled Caste and "Other communities". "The Indian population shows a sustained high frequency of B gene which attains its peak values in the north-western region". Among the tribes, the highest averages are recorded among the Tibeto-Burman speakers. The O gene seems to be co-related with Dravidian speakers with an exceptionally high frequency of 70% among Dravidian speakers ("Other communities") in Maharashtra. The incidence among the Shompen (ST-Austro-Asiatic) is 90%, and "ethnic differences are indicated by the relatively low average incidence of O gene in the tribal populations ... in almost all States". While the M gene under the MN system generally shows a north-south gradient with high averages in the north, the *Vēttuvan*, a Scheduled Caste community of Kerala shows an odd high value of 73.12%.

Maps 108 & 108A exhibit the incidence of the R₁ Haplotype under the Rh system, and 109 & 109A of the r Haplotype; 110 & 110A, of the FY^a gene under the Duffy system, 111 & 11A of the Hp¹ Haptoglobin gene, 112 & 112A of the G-6PD deficiency, 113 & 113A of the PGM¹ gene, 114 & 114A of the ESD² gene and 115 & 115A of the AK¹ gene, the maps numbered 'A' showing the particulars for the Scheduled Tribe communities. Generally, all these traits show more of a geographical rather than a linguistic co-relation. Language certainly does not seem to be a *determinant* of a genetic trait, though the geographic distribution of the language families tend to lump certain traits together among the speakers of the respective languages.

It is hoped that a more unbiased version will perhaps be prepared, as this *Atlas* will be a reference document for several scholars, and for several years into the future. When preparing a revised version, it is also suggested that several of the pie diagrams and bar charts which require so much of explanation may be replaced by suitable colour coding.

Bibliography

- Singh, K.S. 1992. *People of India: An Introduction*. Vol. I. People of India.
- Singh, K.S. & Manoharan. 1993. *Languages & Scripts*. Vol. IX. People of India. Oxford University Press. Delhi, etc.
- Walker, Anthony R. 1986. *The Toda of South India, a New Look*. Hindustan Publishing Corporation. Delhi.

T. MADHAVA MENON

COMMENTS ON THE REVIEW

We have discussed the notion of the ethnographic community at great length in the chapter on "Conceptual Framework and Methodology" in the first volume, *People of India: An Introduction* in which a community has been defined in terms of endogamy, occupation or a set of occupations and perceptions. Going by this definition and considering the stratification sharply emerging within a community, there is a diversity of perceptions within it. There inevitably is an overlap in traits. This has been amply reflected in the data generated by the People of India (POI) project and the Atlas.

We have tried to be explicit with our concepts as far as possible and those familiar with cartography will not find any difficulty in understanding them. Language was mentioned as a category in the biological maps because our biological data could be organised better in terms of the

linguistic categories rather than a geographical region. Placing third before the second is an error.

We have followed the general definition of terrain or climatic zones, taking into account, in a very general way, people's own perception of their habitation. The botanical name of some of the local cereals and pulses have been given in the last volume of the national survey which are all of a piece and have to be read together.

Our data for Kerala does show a very high level of consumption of roots and tubers by the SCs and most of the STs. Which of the ST groups do not consume roots and tubers will be looked into closely at the state level when we analyse our data. Even otherwise, it is within the permissible limits of errors.

The map on polyandry deals with fraternal type of polyandry. Our investigators have reported that fraternal polyandry has not been entirely given up among the Todas. While the institution of polyandry has disappeared, the observance of the "ritual" has persisted as a cultural survival. We are not concerned with prevalence of polyandry in the past. We have reported the current situation. The non-plotting of the non-fraternal type of polyandry was an omission.

As mentioned earlier, there is bound to be an overlap in reporting occupation or any other trait within a community. The data on animal husbandry relates to the main occupation of animal husbandry.

Our data on traditional health system and modern medicare show preference for the two systems among many communities all over India including Kerala. We have gone strictly by the response of the communities to our four questions. As it is known, the urban groups show a greater preference for modern medicare than the rural and tribal folk, who depend mostly on the traditional system.

Map 72 on languages was initially prepared by our senior linguist Dr. Trivedi and was later scrutinised and revised by us. It was not based on the POI data. It deals with major languages-not all languages- and distribution of language families over larger areas of their concentration. Our note clearly states that the tribal languages belonging to the Dravidian language family are distributed in pockets in central and eastern India. They have been shown by number, not area, for reason of space, as has been done in the case of the non-Dravidian and smaller Dravidian languages located in the Dravidian language family area. Entry 39 on Kurukh stands within the Austro-Asiatic area, as the Munda and Oraon populations live together in Ranchi district. Entry 52 which should actually read 59 -and vice-versa-has been wrongly placed within the Austric (Korku) area. Kui has been shown by the number. This is also true of entry

33 which should have been placed a little southward. In the following map (73) based on our firsthand data, we have shown the Dravidian languages spoken in Madhya Pradesh, Maharashtra, Orissa, West Bengal and other states, even though it is not possible to mention all languages including the pidgins and all language speakers in an atlas. This has been done in a separate volume on languages.

It should be noted that we have shown the distribution of languages as it exists today. We are not concerned with history. We have used the notion of historical immigrants strictly in the context of the immigrants from the Maldives Island, or the perception of the communities like Banjari, Marathi, Lambadi etc. whose movement could be traced to the medieval and colonial periods.

In our project we have particularly studied the immigrants from various language areas scattered over the country. The Bhil, according to the present view belongs to Indo-Aryan language family, no matter what its historical antecedents were. As a researcher, I have always held the view that the influence of Dravidian languages is far more widespread in the north and north-west than is suggested by the existence of the pockets of tribal languages belonging to the Dravidian family.

As far as map 75 is concerned, according to article 343 of the Constitution of India, the official language of the Union shall be Hindi in Devanagari script and that English language shall continue to be used for all the official purposes of the Union. According to the POI data Hindi and its variants are spoken by the largest number of communities spread over many states. It was in this limited sense that the term Hindi as a "national language" was used, without implying any disrespect to the more ancient and richer languages.

The north-eastern region is linguistically most heterogeneous, because of the presence of so many tribal languages. A reason for the lowest incidence of monolingualism in Bihar is that the communities in the state have reported speaking various languages like Maithili, Bhojpuri, Magadhi, tribal languages etc., at home and outside.

The map on scripts (76) shows only two scripts in use in Lakshadweep namely Malayalam and Thana, as reported by investigators. We have strictly gone by the response of the people to our question on the use of the scripts at home and outside.

We have not shown any ubiquitous spread of Indo-European languages in South India. As we have explained such languages are spoken by a small number of communities.

As we have mentioned in our volume on languages, there is no one to one correspondence between language and "culture". There is no relationship between a language and a genetic trait. It was however easier for us to organise our biological data for the Scheduled tribes and other categories in terms of language family rather than geographical distribution. The linguistic and geographic categories converge in many states, if not all.

Colour coding was estimated to be too expensive for the project.

K.S. SINGH

REPLY TO THE COMMENTS

T. MADHAVA MENON

The draft of the *Review* had been sent to Dr K.S. Singh for his comments, published above. We appreciate the promptitude with which he responded as well as for his intellectual candour admitting at least the more obvious mistakes. The reviewer replies to Dr Singh's comments as follow.

"The (People of India) project identified various units of the Indian community and the controversy about 'community' was as irrelevant as that about the definition of 'tribe'. No definition can be too precise at any point of time. Definitions change." so said Dr Singh in the First Volume of the POI series (1992:28-29). He has now reinforced this, stating that the "definition" in the First Volume leads to a "diversity of perceptions within it. There inevitably is an overlap in traits". Thus, the "definition" is neither exclusive and nor "definite"!

The problem of understanding the concepts explained in the "Introduction" to the *Atlas* was not in the cartography, but in the language. Dr Singh has admitted that placing "third" before the "second" was an error. The bars of 1 mm height show not the "value of a community", whatever that may mean, but the number of communities concerned. Usages like the "value of a community", when it is the number that is involved, should have been avoided.

The physiographic features and climatic parameters have objective units of measurement. When, in this *Atlas*, a departure is made from this cartographic convention, a note of explanation would have been in place, but is lacking. The explanation now given is too "general". If it was inconvenient to include the botanical names of cereals and pulses in the *Atlas*, a cross reference could have been given to the last volume of the

national survey, which incidentally has not yet been available to us. Normally, a map should tell its own story.

Dr Singh admitting that the "non-plotting of the non-fraternal type of polyandry was an omission", asserts that the fraternal type has not been entirely given up among the Todas because "While the institution of polyandry has disappeared, the observance of the 'ritual' has persisted as a cultural survival". Perhaps he has in mind the ceremony of the symbolic bow and arrow as a token of social acceptance of paternity. If this is considered equivalent to persistence of polyandry, it goes against his repeated assertion that he was "not concerned with history".

The explanation, offered now, of the "mysteries" of Map 72 on the distribution of the "Major Languages", makes it "curioser and curioser". A map has, primarily, to co-relate the geographical location with the feature concerned. This is not true of several languages, eg., Gondi, spoken by over a million people. The map shows it as limited to Andhra Pradesh; further comment seems to be needless. Dr Singh admits that entry 52 should read as 59 and vice versa. Even if "smaller Dravidian languages" are designated by numbers, a map should show them approximately at the places where the language occurs.

The problem in Map 73 is the note contrasting Dravidian languages in the north as due to "recent" migration, whereas Indo Aryan languages in the south are due to "historic" migration. But, says Dr Singh, he was not "concerned with history"!

I am grateful to Dr Singh for pointing out the subtle difference in Article 343 of the Constitution between Hindi, the "official language" of the Union, and English which shall continue to be used for all the official purposes of the Union. The description of Hindi as the "national language" in map 75 was certainly inexact.

"The script used (by the Koya of Lakshadweep) is Arabic, whereas the language and script with others are Malayalam" (p 63); this is true of Malmi (p 71) and of the Melacheri (pp 77) - so say Singh *et al* in Vol. XXVII of the People of India series *Lakshadweep*, 1993. It is the communities of the Minicoy island of this group who use Thana and Malayalam scripts.

The correlation between language groups and biological traits does not make language "a category" in biological traits unless it is another example of extreme liberty with language!

PUBLICATIONS RECEIVED

Takali Bibliography. Dr.A. Razaludeen, Publication Dept. Kerala Unty. Gift from Dr. Ravisankar S. Nair

A Dictionary of Dialectical Bengali Language. Vol.1. Pratap Kumar Khosh, Education Dept. Govt. of West Bengal. Gift from the Higher Education Dept. of West Bengal.

The Hittities. O.R. Gurney. Pelicon. Original Gift from Subhadra Kumar Sen.

Chenchus. P.K. Bhowmick, Institute of Social Research and Anthropology, Calcutta. Gift from Prof. V.I. Subramoniam.

Apna Aadar's meem hum. Dr. B.D. Jetty, Pitamber Publication Company, Delhi. Gift from Dr. B.D. Jetty.

Central Dravidian Comparative Morphology. B.Suvarchala, Navarang Publication, New Delhi. Gift from Prof. V.I. Subramoniam

Manda Vocabulary. Dr. B. Ramakrishna Reddy, Dept. of Osmania University Publications, Hyderabad. Gift from Prof. V.I. Subramoniam

Hegemony of English. Hans R. Dua, Yesodha Publications, Mysore. Gift from Prof. V.I. Subramoniam

Cholanaikans of Kerala. B.Ananda Banu, A.S.I. Calcutta, Gift from Prof. V.I. Subramoniam

Tulu Lexicon, Vol.II. Dr. U.P. Upadhyaya, Govinda Pai Research Centre, Manipal. Gift from Prof. U.P. Upadhyaya.

Orumaippāṭirkku Vinṇeli. Nellai Cu. Muttu, Vijayapatippaagam, Coimbatore. Gift from Prof. V.I. Subramoniam

Profiles in Indian Languages. Arun Kumar Viswas, Indian Institute of Technology, Gift from Prof. V.I. Subramoniam

Dravidian Studies. Selected Papers. M.B. Emeneau, Motilal Banarsddas publications, New Delhi, Gift from Prof. V.I. Subramoniam

C.I.I.L Publications

A programmed Course in Tamil. G. Sankaranarayanan.

Kannada Speech sounds for Second Language Learners. K. Nanyan Dev Araiah

Frame work for Testing Kannada Reading. G.P. Purushottama, 1994.

Darsana Bharati, A Sanskrit Reader for P.G. Students. K.K. Misra and N. Radhakrishna Bhatt, 1994.

A study of Telugu Regional and Social Dialects. J. Venkateswara Sastri 1994.

Phonemic and Morphophonemic Frequency Count in Oriya. Asok R. Kelker 1994.

Konyak Folk Literature. K.S. Nagaraja, 1994

Structural Analysis of Kashmiri Folk Tales. Mrs. Lalitha Handoo.

Konyak Hindi-English Dictionary. K.S. Nagaraja, 1994.

A Linguistic Study of the Problems in Learning Tamil by Tribal Childrem of Nilagiri. Sivakumar.

Bilingualism in a multi-lingual Society. Ajit K.Mohanthy.

A Linguistic Study of Errors. R. Nanyariah

Technology and Languages. B.B. Rajapurohit

Spatial Aspects of Language. N.H. Itagi

Mao Naga Grammar. P.P. Giridha. Prof. V.I. Subramoniam presented all these books to the library.

Papers Received by the I.S.D.L. Library

Development of forms and functions interrogatives in children: A longitudinal Study in Tamil. R. Vaidyananthan, Nair Hospital, Bombay, Child languages 15 (1988)

Mother-tongue and the brain: A Japanese Specialist's surprising discovery. Howsand Bralyn.

Is your brain really necessary. John Lorber, a British Neurologist.

Math and Sex: An girls born with less ability. A Johns Hopkins Group.

Development of Noun Inflections of Verb Paradigms: A longitudinal study in Tamil. Dr. Vaidyanathan.R.

Dravidian and Mangolian 'to be,/to become, to exist(in a place)' and the negative verbs of existence. Jaroslav Vacek, Praha.

Teaching Tamil in the 1970 and 1980s. Jaroslav Vacek.

Multiple-Authorship Pattern in Linguistics, B.A. Sarada, S.R.L.C. Mysore.

Citation Analysis of the Journal Indian Linguistics, 1971- 1980. B.A. Sarada.

Language Power and Idiology: Issues in Applied Linguistic Research. Yamuna Kachru, University of Illinois.

Linguistic Curriculum: A need for Change. B. Lakshmi Bai, Osmania University, Hyderabad.

The papers presented in the Seminar conducted on the Silver Jubilee Celebration day of C.I.I.L.

Ancient Migration and Languages. Prof.V.I. Subramoniam, Foundation Day Features.

Papers on Goals and Strategies for Development of Indian Languages.

NOW AVAILABLE

Dravidian Encyclopaedia - Volume I

The focus of this volume is on the political history of the States south of the Vindhyas. The pre-history, the findings in the excavations at Mohenjo-daro, Harappa etc., are completely covered. The source material for history like the Vedas, the Sangam literature, epigraphs etc., have been fully discussed. The contribution to ayurveda, siddha and military sciences are covered with facts. The developments in chemistry, physics and mathematics have also been traced in this volume.

The Dravidian Encyclopaedia adopts:

- *A balanced presentation where there are two opinions on an academic problem.*
- *An objective presentation without any attempt for glorification.*
- *International co-operation in securing articles for the Encyclopaedia.*

It summarises the findings of 1100 published works and 3500 journal articles. Photo-composed and printed in offset. 725 pages of demi 1/4 size, 80 pictures. Get-up of the Encyclopaedia- Britannica.

Price : Rs.500/-

Dravidian Encyclopaedia - Volume II

Total No. of Articles 987 Contributed Articles 142

Written by Staff 845 Line Drawings 6

Maps 4 Photographs 61

Tables 36 Number of pages 841 + 8 + 7

Index 60 Total 916 pages.

Well-printed and documented

BOOK REVIEWS

Two copies of the books to be reviewed should be sent. Twentyfive off-prints will be given to the reviewer and ten to the publisher.

SUBSCRIPTION

The annual subscription from 1st January to 31st December each year is Rs. 200 for India, US \$ 40 for countries other than India. Single number Rs. 110 or US \$ 20 each.

The members of the Dravidian Linguistics Association are eligible to get a free copy of *ijdl* and DLA News, each from the Treasurer of the DLA.

Accredited book sellers are eligible to get 10% commission for the journal.

ADVERTISEMENT

Only from publishers and academic institutions advertisements are accepted. One insertion, full page Rs.300 or U.S.\$60: For six insertions, full page Rs.250 or U.S.\$50 each.

CORRESPONDENCE

All matters relating to subscription, sales, advertisements and change of address of the subscribers, manuscripts and books for review should be addressed to the Chief Editor, *ijdl*, St. Xavier's College P.O., Trivandrum, S. India.

Please note that the cheques drawn may be in favour of the Chief Editor, *ijdl*, St. Xavier's College P.O., Trivandrum, S.India.